

# The Pricing Kernel Puzzle: Survey and Outlook

by

Horatio Cuesdeanu

and

Jens Carsten Jackwerth

We update the survey of the pricing kernel puzzle of Jackwerth (2004). We document the pricing kernel puzzle in several markets and present the latest evidence concerning its (non-)existence. Econometric studies are detailed which test for the pricing kernel puzzle. Explanations of the pricing kernel puzzle are discussed and we suggest areas where more research is needed.

Key words: pricing kernel puzzle, stochastic discount factor, options, S&P 500

JEL classification: G12, G13

Horatio Cuesdeanu is from the University of Konstanz, PO Box 134, 78457 Konstanz, Germany, Tel.: +49-(0)7531-88-4326, Fax: +49-(0)7531-88-3120, [horatio.cuesdeanu@uni-konstanz.de](mailto:horatio.cuesdeanu@uni-konstanz.de)

Corresponding author: Jens Jackwerth is from the University of Konstanz, PO Box 134, 78457 Konstanz, Germany, Tel.: +49-(0)7531-88-2196, Fax: +49-(0)7531-88-3120, [jens.jackwerth@uni-konstanz.de](mailto:jens.jackwerth@uni-konstanz.de)

## Contents

<b>Preface .....</b>	<b>4</b>
<b>1. Introduction and a simple model of the pricing kernel .....</b>	<b>5</b>
<b>2. The pricing kernel puzzle.....</b>	<b>8</b>
<b>2.1. The initial studies .....</b>	<b>10</b>
<b>2.2. Overly restricted pricing kernels .....</b>	<b>13</b>
2.2.1. Assumptions restricting the functional form of the pricing kernel: stochastic volatility, stochastic jump models .....	14
2.2.2. Assumptions restricting the functional form of the pricing kernel: power and exponential utility functions .....	15
2.2.3. Assumptions restricting the functional form of the pricing kernel: the maximum entropy approach .....	17
2.2.4. Assumptions restricting shape of the risk-neutral distribution.....	17
2.2.5. Ross (2015) recovery .....	18
<b>2.3. The pricing kernel puzzle and option returns .....</b>	<b>20</b>
<b>2.4. Applications .....</b>	<b>22</b>
<b>2.5. The pricing kernel puzzle in exclusively forward-looking data .....</b>	<b>24</b>
<b>3. Empirical Studies.....</b>	<b>25</b>
<b>3.1. The S&amp;P 500 index market.....</b>	<b>26</b>
<b>3.2. The German DAX 30 index market .....</b>	<b>28</b>
<b>3.3. The British FTSE 100 index market .....</b>	<b>30</b>
<b>3.4. Other index markets.....</b>	<b>30</b>
<b>3.5. Non-index asset markets .....</b>	<b>32</b>
<b>3.6 The shape of the empirical pricing kernel .....</b>	<b>35</b>
<b>4. Econometric assessment of the pricing kernel puzzle.....</b>	<b>36</b>

<b>4.1. Sample variation bounds .....</b>	<b>36</b>
<b>4.2. Perturbations of options data .....</b>	<b>37</b>
<b>4.3. Statistical tests .....</b>	<b>38</b>
<b>5. Solutions .....</b>	<b>41</b>
<b>5.1. Models with a single state variable.....</b>	<b>42</b>
5.1.1. Heterogeneous investor models .....	42
5.2.2. Piece-wise approximations of the pricing kernel .....	46
<b>5.2. Models with several state variables.....</b>	<b>47</b>
5.2.1. Models with volatility as a second state variable.....	47
5.2.2. Regime-switching models.....	49
<b>5.3. Behavioral and sentiment models .....</b>	<b>51</b>
5.3.1. Demand based models .....	51
5.3.2. Models with probability weighting functions.....	53
<b>5.4. Ambiguity aversion models .....</b>	<b>54</b>
5.4.1. The theoretical pricing kernel under ambiguity aversion .....	55
5.4.2. The pricing kernel puzzle in a model of ambiguity aversion .....	57
<b>6. Bounds on option prices.....</b>	<b>59</b>
<b>7. Conclusion and outlook.....</b>	<b>61</b>
<b>Bibliography .....</b>	<b>63</b>

## Preface

It has been a while since the literature on the pricing kernel puzzle was summarized in Jackwerth (2004). That older survey still covered the topic of risk-neutral distributions, which was itself already surveyed in Jackwerth (1999). Much has happened in those years and estimation of risk-neutral distributions has moved from new and exciting in the last half of the 1990s to becoming a well-understood technology. Thus, the present survey will focus on the pricing kernel puzzle, which was first documented around 2000 and first surveyed in Jackwerth (2004). The present work adds much breadth in terms of international evidence, statistical tests, and economic explanations of the puzzle. New challenges in the field are described in the process.

## 1. Introduction and a simple model of the pricing kernel

The pricing kernel  $m$  is of fundamental concern to all of modern finance as it is the basis for all pricing:

$$E[mR] = 1 \quad (1)$$

where  $E[\cdot]$  is the expectation under the physical (true) probabilities  $p$  across states and  $R$  is the return in each state. The pricing kernel  $m$  is the ratio of state prices  $\pi$  and physical probabilities  $p$  or, alternatively, of discounted risk-neutral probabilities ( $q/R_f$ ) and physical probabilities:

$$m = \frac{\pi}{p} = \frac{q}{R_f p} \quad (2)$$

The pricing kernel informs us on how we need to adjust payoffs  $X$  such that we can take simple expectations in order to obtain the price of the security. It thus contains important information about the investor's assessment of different states: payoffs in states associated with low wealth/consumption are valued highly ( $m$  is large).

We can appreciate the link between the pricing kernel and preferences in a simple one-period representative investor economy with concave utility of end of period wealth  $U(R)$ . The representative investor maximizes end of period utility while investing all his unit wealth:

$$\begin{aligned} \max_{R_i} E[U(R_i)] &= \sum_{i=1}^N p_i U(R_i) \\ \text{s.t. } E^*[R_i] / R_f &= \left( \sum_{i=1}^N q_i R_i \right) / R_f = 1 \end{aligned} \quad (3)$$

where  $R_i$  are the cum-dividend aggregated market returns across the  $i = 1, \dots, N$  states,  $p_i$  are the physical probabilities,  $U$  is the utility function,  $q_i$  are the risk-neutral probabilities. One plus the interest rate over the period is  $R_f$ . Introducing the Lagrange multiplier  $\lambda$ , we write the  $N$  first order conditions:

$$p_i U'(R_i) - \lambda q_i / R_f = 0 \quad \text{for } i = 1, \dots, N \quad (4)$$

and solve for  $\lambda$  after summing the  $N$  equations (4):

$$R_f \left( \sum_{i=1}^N p_i U'(R_i) \right) = \lambda \quad (5)$$

We can now find the pricing kernel in state by state by inserting the explicit solution for the Lagrange multiplier into the first order condition (4):

$$m_i = \frac{q_i}{R_f p_i} = \frac{U'(R_i)}{R_f \sum_{i=1}^N p_i U'(R_i)} \quad \text{for } i = 1, \dots, N. \quad (6)$$

Equation (6) informs us that the pricing kernel is proportional to marginal utility. Any insight into the pricing kernel thus translates into knowledge about investor preferences in our economy. In particular, standard concave utility functions, such as power and exponential utility, lead to positive and monotonically decreasing pricing kernels.<sup>1</sup>

Alternatively, we can also express absolute risk aversion directly:

$$ARA_i = -\frac{U''(R_i)}{U'(R_i)} = \frac{p'}{p} - \frac{q'}{q} \quad (7)$$

which is a positive function as long as the utility function is concave and marginal utility is positive.

Such insights into investor preferences are particularly useful since utility functions are notoriously difficult to estimate. One way of eliciting preferences from investors is to survey the investors. Overberg (2009) and Andreoni and Sprenger (2012a) collect a number of such

---

<sup>1</sup> There exists a close connection to the concept of stochastic dominance which expresses dominance relations between probability distributions on which all investors of a certain class agree. Our setting of positive and decreasing pricing kernels uses the class of risk-averse investors (i.e., those with concave utility functions) and the corresponding concept is the one of second order stochastic dominance.

surveys, which try to estimate investor risk aversion. Barsky, Juster, Kimball, and Shapiro (1997) is an example of a survey which elicits risk tolerance, time preference, and intertemporal substitution. Wang, Rieger, and Hens (2013) have an international survey covering 52 countries covering time preferences and loss aversion.

A second line of research uses market data to infer the (parametric) utility function of a representative investor. A starting point is the equity premium puzzle of Mehra and Prescott (1985). Here, a stylized economy with a representative investor economy with power utility is being calibrated to market data. The resulting risk aversion coefficients tend to be much too high when compared to survey based estimates. This literature has been continued in Kocherlakota (1996) and Mehra (2006), with international evidence added in Pozzi, de Vries, and Zenhorst (2010). Closely related is the work by Bartunek and Chowdhury (1997) who use power utility and Benth, Groth, and Lindberg (2010) who use exponential utility instead; both papers calibrate to options data. While the equity premium literature equates the forward looking physical probability distribution with the historical distribution and determines the intertemporal rate of substitution solely through the risk aversion coefficient, Andersen, Fountain, Harrison, and Rutstroem (2014) suggest to elicit physical probabilities, too, and Andreoni and Sprenger (2012a, 2012b) and Laury, McInnes, Swarthout, and Von Nessen (2012) additionally estimate time preferences.<sup>2</sup>

---

<sup>2</sup> An interesting observation reconciling the two estimation methods (surveys/experiments vs. market based) can be found in Haug, Hens, and Woehrmann (2013) who argue that the typical inclusion of background risk in market studies leads to larger risk aversion coefficients than in surveys or experiments which tend to ignore background risk. Related is the work by Meyer and Meyer (2005a, 2005b) who point out that the risk aversion estimate also depends on which exact quantity (consumption or wealth) enters the (direct or indirect) utility function in which numerical form (say, return or rate of return).

## 2. The pricing kernel puzzle

We are more interested in a third approach, the direct estimation of the pricing kernel  $m$  via Equation (2), which is possible and involves estimation of (one plus) the riskfree interest rate  $R_f$ , of the physical probabilities  $p$ , and of the risk-neutral probabilities  $q$ . Estimation of the interest rate is an easy task as the discounting effect is small over the typical horizons of 30 to 60 days. The methods of choice to obtain the physical probabilities  $p$  are kernel density estimation of past index returns on the S&P 500 (e.g. Jackwerth (2000), Ait-Sahalia and Lo (2000)) or the estimation of GARCH models (e.g. Rosenberg and Engle (2002), Barone-Adesi, Engle, and Mancini (2008), Barone-Adesi and Dall'O (2010)). Estimation of the risk-neutral distribution is by now a well-established field of research and a large literature covers it; see the surveys by Jackwerth (1999), Jackwerth (2004), and Bahra (1997). Given a large enough cross section (more than 10 option strike prices), most methods perform relatively similar and yield the desired risk-neutral distributions where one typically uses the SPX options on the S&P 500 index with typical maturities of 30 to 60 days. Most studies work with the S&P 500 as it is seen as a reasonable proxy for the market return even though it does not cover all investment opportunities of a representative investor. Moreover, many liquid options written on the S&P 500 index exists and, hence, one is able to back out risk neutral distributions with less estimation error than other, less liquid indices. In alternative models of the economy, the pricing kernel is a function of aggregate consumption or total wealth, and, once again, the S&P 500 index is seen as a good proxy for these variables.

The most popular methods for obtaining the risk-neutral distribution are probably the mixture of two lognormals or the spline based approach. Mixtures of three lognormals were introduced by Melick and Thomas (1997) and then reduced to two lognormals in Bahra (1997). The original spline approach was introduced by Shimko (1993) and refined in Bliss and Panigirtzoglou (2002, 2004). A particularly fast and stable method can be found in

Jackwerth (2004). A recent careful but also quite complicated implementation is Figlewski (2010). In order to model the implied volatility surface and the corresponding risk-neutral distribution for multiple horizons, Fengler and Hin (2015) use a semi-parametric approach based on B-splines while Ludwig (2015) suggests the use of neural networks.

Option-implied distributions are frequently used to assess market expectations, see the above surveys and the recent applications by the European Central Bank in EZB (2011) or David and Veronesi (2011) who use option-implied measures to predict macro variables.

In Figure 1, we find representative risk-neutral distributions for the S&P 500 in the US, the DAX 30 in Germany, the FTSE 100 in the UK, and the Nikkei 225 in Japan. The physical distributions are estimated using a simple kernel density estimator based on historical returns. More details can be found in Jackwerth (2004) from where the examples stem.

[Figure 1 about here]

After dividing the risk-neutral distribution by the physical distribution, we obtain the empirical pricing kernels, which are depicted in Figure 2. Note the tilde-shaped hump around at-the-money, which is inconsistent with Equation (4) according to which the empirical pricing kernel is monotonically decreasing in returns since it is proportional to the marginal utility of a risk-averse investor. For such risk-averse investor, utility is concave and marginal utility is decreasing. Moreover, equilibrium is ruled out as a non-decreasing pricing kernel implies the existence of a portfolio that stochastically dominates the market, see Sections 5 and 6. A non-decreasing pricing kernel hence clashes with our basic intuitions and contradicts most standard market models. The violation of monotonicity has been labeled as the “pricing kernel puzzle” and we will investigate possible explanations in Section 5.

[Figure 2 about here]

## 2.1. The initial studies

The birth of the pricing kernel puzzle can be determined pretty accurately in the year 1996. For some time by then there was some general unease about the consistency of time-series parameters of the physical distribution stemming from observed returns and the risk-neutral parameters implied in option prices written on the same underlying. Bates (1996a, 1996b) points out that the two stochastic processes seem to be incompatible. Moreover, the time was ready to compare the physical distribution to the option-implied risk-neutral distribution. The physical distribution could be obtained either by kernel density estimation or simply through calibration of a parametric return process to the historical returns, say a GARCH(1,1). Also, Rubinstein's (1994) seminal article allowed for the first time to recover risk-neutral, option-implied distributions.<sup>3</sup> Jackwerth and Rubinstein (1996) extended and applied that technique to the S&P 500 index options. Taking the last step of finding the empirical pricing kernel through dividing the risk-neutral probability distribution by the physical distribution seems obvious in retrospect but was not quite so clear at the time.

The first working paper draft of Jackwerth (2000) was dated March 1996 and actually suggested to approximate the more complicated risk aversion functions (Equation 7) as opposed to the more straightforward pricing kernel (Equation 2) directly. S&P 500 index options from 1986 through 1995 are filtered and used to obtain 31-day implied risk-neutral distributions which are smoothest while pricing almost all observed options in the monthly cross section within bid/ask spreads. The physical distribution is based on 31-day, non-overlapping returns within moving, 4-year historical windows. The returns are then smoothed through a kernel density estimator. A number of robustness checks confirm the result that the empirical risk aversion functions are u-shaped and negative around at-the-money during the

---

<sup>3</sup> The earlier theoretical approach of Breeden and Litzenberger (1978) suffers from numerical implementation problems.

post-87-crash period while they are mainly positive and decreasing during the pre-crash period. In particular, replacing the physical distribution with a simulated conditional distribution from a fitted GARCH(1,1) model only strengthens the results. A peso problem also cannot explain the pricing kernel puzzle, since for the first 4 years past the crash, the crash is “visible” in the physical distribution based on the historical returns. Still, the results do not change compared with periods where the crash is no longer visible because, on the date of the investigation, the crash lies more than 4 years into the past.<sup>4</sup> A related paper is Figlewski and Malik (2014) who use a lognormal distribution for the physical distribution, fix the mean with a 5% risk premium, and use a volatility, which is partly historical and partly based on VIX. Using data on an exchange traded fund replicating the S&P 500, they confirm the pricing kernel puzzle.<sup>5</sup>

A little later, Ait-Sahalia and Lo (1998) suggested a method for backing out risk-neutral, option-implied distributions based on applying a kernel regression to the implied volatilities of observed options. The obtained implied volatility function can then be translated into a function of call prices and, with the help of Breeden and Litzenberger (1978), they obtain the risk-neutral distribution as the second derivative of the scaled call option price function. The authors derived the pricing kernel independently of Jackwerth (2000) as the ratio of the risk-neutral distribution and the physical distribution obtained through a kernel based estimator. Based on half-yearly returns during the year 1993 they can document the pricing kernel

---

<sup>4</sup> In a richer setting, Ziegler (2007) confirms the point that a Peso problem cannot explain the pricing kernel puzzle. See Section 5.1.1 for details.

<sup>5</sup> See also Hill (2010) for separately fitting the risk-neutral and the physical distribution to variance-gamma and normal-inverse-Gaussian processes. As he does not economically restrict the parameters between the two measures, the resulting distributions can differ from each other and their ratio exhibits the pricing kernel puzzle.

puzzle. This research was published in Ait-Sahalia and Lo (2000) where the authors very graciously delayed publication so that their paper would not appear in print before the publication of Jackwerth (2000) which was started earlier but was long delayed at the journal.<sup>6</sup>

The third of the canonical models, which is typically cited to establish the pricing kernel puzzle, is Rosenberg and Engle (2002). Using monthly data from 1991 to 1995 on the S&P 500 index options, they start by obtaining the physical distribution from the parametric GARCH model of Glosten, Jagannathan, and Runkle (1993) fitted to historical returns. They next specify the pricing kernel parametrically, which allows them to obtain the risk-neutral distribution and thus derive model-implied option prices.<sup>7</sup> The parameters of the pricing kernel are optimized such that the sum of squared option pricing errors is being minimized. First, a monotonically decreasing pricing kernel is being fitted but mispricing can be much reduced when more flexible functional forms for the pricing kernel are allowed, leading to the pricing kernel puzzle yet again. A closely related minor extension is Yang (2009) and another one, using B-splines for the pricing kernel, is Audrino and Meier (2012).

---

<sup>6</sup> Belomestny, Ma, and Haerdle (2014) is a modern implementation of Ait-Sahalia and Lo (2000) and also uses kernel estimation for the physical and the risk-neutral distributions.

<sup>7</sup> Dittmar (2002) suggests an alternative parametric specification of the pricing kernel based on a Taylor series expansion in terms of returns on wealth. He restricts the signs economically ( premia should be positive for the equity and skewness risk premium, negative for volatility and kurtosis; compare also Schneider and Trojani (2015) for a similar idea in Section 2.2.5). Empirically, he finds some evidence of a u-shaped pricing kernel, which is confirmed by Schweri (2010). For a related Taylor series approximation of the pricing kernel, see Chabi-Yo (2012).

A final interesting and early paper is Carr, Geman, Madan, and Yor (2002). They fit a specific Levy process, the CGMY process, to historical returns and option prices. The CGMY process is more flexible when compared to standard diffusion processes due to the inclusion of jump components. The resulting risk-neutral distribution is fatter-tailed than the physical distribution and the pricing kernel is u-shaped. This constitutes a pricing kernel puzzle as well but of a different kind than the so far discussed tilde-shaped pattern which occurs around at-the-money (see Section 3.6 for a discussion on tilde-shaped versus u-shaped pricing kernels). Surprisingly, the authors state for this particular stochastic process that the risk-neutral parameters can be chosen independently of the physical parameters. Thus, the stochastic process does not restrict the pricing kernel beyond positivity, i.e. the no-arbitrage condition. In contrast, for diffusion processes the risk-neutral and the physical parameters are linked through the presence of risk premia. Wu (2006) extends the CGMY process by allowing for asymmetric distributions. Moreover, he explicitly characterizes the pricing kernel in his model as the multiplication of two exponential components, which depend on the market price of positive jump risk multiplied by the positive jumps and a similar expression for the market price of negative jump risk times the negative jumps where the two prices of risk can differ. While he is motivated by the pricing kernel puzzle which he depicts in his figure 4, he unfortunately does not plot the empirical pricing kernel based on his estimation using S&P 500 returns and option prices.

## **2.2. Overly restricted pricing kernels**

The canonical models did not give much thought to the direction of their inference as it seemed rather natural to back out pricing kernels, which tend to be difficult to estimate, from risk-neutral distributions, which tend to be easy to obtain from cross sections of option prices (see the argument in Jackwerth (2004, p. 15)), and physical distributions, which can be estimated from historical data, albeit with estimation error.

### **2.2.1. Assumptions restricting the functional form of the pricing kernel: stochastic volatility, stochastic jump models**

While the canonical early papers backed out the pricing kernel, other researchers tried to find the forward looking physical probabilities by assuming a functional form for the pricing kernel. However, imposing severe restrictions on the pricing kernel can lead to estimates which will than no longer exhibit the pricing kernel puzzle despite its presence in the data. E.g., Chernov and Ghysels (2000) fitted the Heston (1993) model to S&P500 index returns and option prices. The Heston (1993) model has a market price of risk component and a further volatility risk component where the latter depends on the long run volatility level and the speed of mean reversion of the volatility process. Chernov and Ghysels (2000) argue that the simultaneous estimation of the physical and the risk-neutral processes can be beneficial in select circumstances. While the paper provides expressions for the pricing kernel, it is not immediately clear that the pricing kernel puzzle can be generated altogether, given the restrictive choice of a few constant risk premia, which account for the parameter differences between the physical and the risk-neutral versions of the model. Similarly, Pan (2002) extends the Bates (2000) model, which in turn is an extension of the Heston (1993) model. Stochastic jumps are being added to the model, as well as stochastic interest rates and stochastic dividend yields, which however play a lesser role. Still, when plotting the pricing kernel using the estimated values of Pan (2002), it emerges to be monotonically decreasing and thus does not exhibit the pricing kernel puzzle. Two more examples of such stochastic volatility, stochastic jump model which cannot generate the pricing kernel puzzle are Bates (2008) and Santa-Clara and Yan (2010). So far, no stochastic-volatility, stochastic-jump model could be fitted to the risk-neutral and physical processes simultaneously, while at the same time

exhibiting the pricing kernel puzzle,<sup>8</sup> but compare the solutions to the pricing kernel puzzle in Section 5.

### **2.2.2. Assumptions restricting the functional form of the pricing kernel: power and exponential utility functions**

A second line of investigation, which specifies the utility function to be of power or exponential type, is also inherently not able to document the pricing kernel puzzle. The leading exponents are Bliss and Panigirtzoglou (2004) who start out with the risk-neutral distribution obtained from option prices. They use the spline method of Bliss and Panigirtzoglou (2002) where a smooth function of implied volatilities is fitted to market implied volatilities. Using the Breeden and Litzenberger (1978) result, the authors obtain the risk-neutral distribution, which they change into the physical distribution through division by the pricing kernel, which is given by the marginal utility of either a power or exponential utility function. They are able to assess the likelihood that the observed, future returns stem from the physical distribution using the method of Diebold, Gunther, and Tay (1998) and Diebold, Tay, and Walis (1999). The idea here is that, under the true distributions, each sampled return can be mapped to a percentile of the cumulative distribution functions. The sample of percentiles then stems from a standard uniform distribution. Given the observed, future returns, one can obtain their percentiles under the derived, physical distribution, and

---

<sup>8</sup> Note that Wu (2006) suggests a pure jump Levy process with different risk premia for negative and positive returns. Fitting very high premia for negative returns, the risk-neutral distribution exhibits a fat left tail and a more normally looking right tail. The physical distribution has two pretty much normally looking tails. These stylized facts line up with our knowledge about the distributions for the S&P 500. The model might thus be able to generate the pricing kernel puzzle, even though, unfortunately, he does not actually depict it in the paper.

compute the likelihood of those percentiles under a standard uniform distribution. The likelihood can now be optimized through the choice of the coefficient of the utility function. The optimal risk aversion coefficient for the power utility turns out to be around 4.

Again, as the parametric functions of the utility functions lead to monotonically decreasing pricing kernels, Bliss and Panigirtzoglou (2004) could not document the pricing kernel puzzle even if it were present in the data. Related is the approach of Duan and Zhang (2014), who assume a power utility formulation to estimate the risk premium from historical returns (using a GARCH model) and options data, while allowing for higher moments to enter the formula for the risk premium. Another early study, which also assumes a power utility, is Weber (2006) in his study of collateralized debt obligations. Finally, Backus, Chernov, and Martin (2011) compare pricing kernels stemming from (real) US consumption to the ones from options on the S&P 500. Finding the pricing kernel puzzle is this time precluded by assuming the Merton (1976) model as the basis for the pricing kernel of the options.

The approach of Bliss and Panigirtzoglou (2004) has been slightly extended in Kang and Kim (2006) who consider additional utility functions but share the above criticism that the parametric choice of utility functions precludes detection of the pricing kernel puzzle in the data. Finally, Benth, Groth, and Lindberg (2010) propose a stochastic volatility model for the physical distribution and an exponential utility function, which they then calibrate to individual option prices by varying the coefficient of the exponential utility for each observed option separately. It is not even clear what it means to have different utility function coefficients at different moneyness levels for the same time-to-expiration. The latest paper to date in this line of thought is Bates (2012) who combines time-changed Levy-processes with (myopic) power utility (see Bates 2006). He estimates the Levy-processes purely based on the time-series of S&P 500 returns and obtains options prices after a change of measure based on fitting the equity risk premium using realized variance. However, options post-crash 1987, out-of-the-money puts, and options during times of high volatility are not well priced.

Conceptually, the same coefficient needs to apply to each cross section. Furthermore, the model is subject to the above criticism of not being able to detect the pricing kernel puzzle by design.

### **2.2.3. Assumptions restricting the functional form of the pricing kernel: the maximum entropy approach**

A third problematic approach is the entropy method of Stutzer (1996). He starts out with the physical distribution as a simple histogram and then finds the maximum entropy risk-neutral distribution, which prices correctly a set of reference assets, namely the stock, the bond, and possibly some known option prices. The resulting risk-neutral distribution is then used to price further derivative assets. However, in the typical scenario of few reference assets the resulting pricing kernel will be monotonically decreasing. Thus, the method will not detect the pricing kernel puzzle, even if it were present in the data, unless a large cross section of option prices is used as reference assets. Extensions of Stutzer (1996) can be found in Alcock and Smith (2014) who implement additional (restrictive) changes of measure based on Haley and Walker (2010).

### **2.2.4. Assumptions restricting shape of the risk-neutral distribution**

Last, Barone-Adesi, Engle, and Mancini (2008) cannot find the pricing kernel puzzle in three years' worth of S&P 500 data (2002-2004) using essentially the methodology of Rosenberg and Engle (2002). However, Barone-Adesi, Mancini, and Shefrin (2013) essentially repeat the study and come to the conclusion, that the pricing kernel puzzle is indeed present in the data. The second paper then goes on to quantify the deviation of the implied pricing kernel from a power utility pricing kernel, using the concept of sentiment, which Shefrin (2008a and 2008b) introduced; see Section 5.1.1. for more details.

The difference in findings is not obvious, given that that the method hardly changed. One reason could be their new way of obtaining the risk-neutral distribution implied in option

prices. Here, they fit a risk-neutral GARCH process where the simulated return distribution should price the observed options as best as possible. However, simulated distributions of daily GARCH models over longer horizons (less than 60 days to maturity to more than 160 days to maturity) tend to be rather normally distributed. Thus, the typical shape of the risk-neutral distribution obtained by other researchers (left-skewed and leptokurtic) might not arise. The implied pricing kernel will then tend to not exhibit the pricing kernel puzzle by design.

Barone-Adesi and Dall'O (2010) change the method for obtaining the risk-neutral distribution yet again and employ a rather crude direct application of Breeden and Litzenberger (1978) by directly using butterfly spreads to approximate the risk-neutral probabilities. Due to the coarseness of the strike prices (\$5 apart in the S&P 500 market), the method introduces jaggedness in the risk-neutral distributions, which is evident in their figures. Here, the poor choice of method for backing out the risk-neutral distributions seems to be the reason for not finding the pricing kernel puzzle in the data.

### **2.2.5. Ross (2015) recovery**

The prevailing thought is that only one of the three quantities, namely risk-neutral probabilities, physical probabilities, and the pricing kernel, can be backed out from the other two. Ross (2015) argues that it would be preferable to use only risk-neutral information, as that is well estimated, and infer both the forward looking physical distribution and the pricing kernel. His insight is that this can be achieved if all risk-neutral transition probabilities are known, as opposed to only the risk-neutral distribution. The difference is that the risk-neutral distribution is one single distribution emanating from the initial (known) state and indicating the (risk-neutral) probability of moving to a future state. The risk-neutral transition probabilities are richer and also indicate the risk-neutral probabilities of moving from all hypothetical initial states to all future states, see Figure 3.

[Figure 3 about here]

While the approach is theoretically very appealing, the determination of the transition probabilities requires some strong assumptions. Ross (2015) assumes that the quarterly transition probabilities are the same in the S&P 500 index options market for the next 3 years and that S&P 500 prices serve as levels. Assuming time-invariant transition probabilities is highly questionable since repeated application of the same transition probabilities leads quickly to fairly normally distributed distributions for longer horizons. However, the risk-neutral probabilities are severely non-normal for short horizons and while they appear more normal over longer horizons does this transition only happen slowly. Furthermore, Ross (2015) needs to assume that the pricing kernel is strictly a ratio of the marginal utility at the future state divided by the marginal utility at the current state. Such pricing kernels are very restrictive, even though they do not preclude the existence of the pricing kernel puzzle.<sup>9</sup>

Jackwerth und Menner (2015) study the empirical implementation of the Ross (2015) recovery and find a number of intractable problems. Obtaining a very smooth implied volatility surface is rather difficult, and working out the matrix of transition distributions leads to oddly shaped transition distributions. Such problems will filter into poorly estimated pricing kernels and physical probability distributions. Jackwerth und Menner (2015) test these physical distributions based on the realized returns, which supposedly stem from them, and

---

<sup>9</sup> Carr and Yu (2012) replace the assumptions on the utility function of a representative investor by assuming that the dynamics of the numeraire portfolio under the physical measure are being driven by a bounded diffusion. Walden extends Ross (2015) recovery to unbounded diffusion processes and Huang and Shaliastovich (2014) to the state dependent, recursive preferences of Epstein and Zin (1989). Schneider and Trojani (2015) suggest recovery based on assumptions on the signs of risk premia on different moments of market returns.

strongly reject the proposed physical distributions<sup>10</sup>, whereas the assumption that physical distributions can be estimated by using historical return distributions cannot be rejected. Audrino, Huitema, and Ludwig (2015) also implement Ross (2015) recovery but use a somewhat different methodology, where their choice of penalizing non-zero state prices leads to u-shaped pricing kernels. They then use the risk-neutral and physical moments to predict profitable trades without investigating their statistical significance. A more obvious first step however seems to be to check, if the realized returns are consistent with the recovered physical distributions.

Jensen, Lando, and Pedersen (2016) develop a recovery framework that makes no assumption of the underlying probability distribution and allows for a closed-form solution. Practical implementation relies only on current option prices for different maturities and hence, there is no need for a full matrix of transition distributions as in the Ross (2015) model. Their empirical results suggest that the recovered statistics have predictive power for future realized returns; although they stress that their empirical implementation primarily has an illustrative purpose. They are able to reject that the full distribution of recovered probabilities exactly matches the true distribution using a Berkowitz test.

### **2.3. The pricing kernel puzzle and option returns**

So far we studied the pricing kernel puzzle in terms of returns of the underlying security, often a broad index such as the S&P 500. But Equation (1), which we repeat here, also holds for option returns:

$$E[m R] = 1 \tag{8}$$

---

<sup>10</sup> A point also made in Borovicka, Hansen, and Scheinkman (2015) who attribute these problems to “misspecified recovery,” which happens when the pricing kernel has non-trivial martingale components.

We start our discussion by looking at the expected return on a call option under the physical measure:

$$E[R_{call}] = \frac{E[(S - K)^+]}{E[m(S - K)^+]} \quad (9)$$

Under the assumption of a monotonically decreasing pricing kernel, call returns should be positive and increasing in moneyness, as, intuitively speaking, the pricing kernel in the denominator shifts mass to the region where the call payoff is zero. A stronger result is presented in Coval and Shumway (2001): the expected return on a call should be greater than the expected return on the underlying, which broadly holds in the data.<sup>11</sup> The authors then investigate returns on option straddles and find evidence of priced volatility risk, which they cannot reconcile with power utility for the representative investor. This evidence is consistent with the pricing kernel puzzle but does not outright prove the case.<sup>12</sup> Broadie, Chernov, and Johannes (2009) caution using unscaled option returns, which tend to be so noisy that one cannot even reject the assumption that the returns were being generated by the Black-Scholes model. Such findings strongly suggest scaling option returns in a suitable way (e.g. straddles as above or by standardizing betas as in Constantinides, Jackwerth, Savov (2013)).<sup>13</sup>

---

<sup>11</sup> Branger, Hansis, and Schlag (2011) do not confirm their result in more recent data, thus documenting the presence of the pricing kernel puzzle in the data. They further argue that stochastic volatility, stochastic jump option pricing models, which also have jumps in the volatility process, can explain those call option returns.

<sup>12</sup> Ni (2009) investigates a test similar to Coval and Shumway (2001) for individual stock options instead of index options. Her surprising results could be due to methodological problems, see the critical discussion in Chaudhury and Schroder (2015).

<sup>13</sup> Constantinides, Jackwerth, Savov (2013) use factor models in order to explain (scaled) option returns. Thus, the pricing kernel then takes a linear form and it is hard to see how the pricing kernel puzzle should arise in such setting.

Chaudhury and Schroder (2015) extend the results of Coval and Shumway (2001) by showing that the pricing kernel is only monotonically decreasing if (conditional) expected returns on certain option positions (called “log-concave” and encompassing long calls, puts, butterfly spreads, and others) increase in the strike price. They confirm the pricing kernel puzzle based on data for the S&P 500 index but fail for individual stock options. This is expected due to the much flatter implied volatility smiles of the individual stock options. Another extension in Bali, Cakici, Chabi-Yo, and Murray (2014) looks at the higher risk-neutral moments of option returns. Song (2012) applies the ideas of Coval and Shumway (2001) to returns on options on volatility in the case of u-shaped pricing kernels.

Bakshi and Madan (2007) present a market model where the pricing kernel is u-shaped since a group of pessimistic investors are shorting the market index. In addition, these investors buy call options as an insurance against a rising index and, hence, are willing to pay a premium for the calls; for more details see Section 5.1.1. Consistent with such market model, Bakshi, Madan, and Panayotov (2010) find evidence for a positive dependence between short-selling activity and expected call returns. Looking again at equation (9), it is clear that a u-shaped pricing kernel directly implies that expected returns of call options with a strike above a certain threshold are negative and decreasing in the strike price. Bakshi, Madan, and Panayotov (2010) document evidence for such a u-shaped pricing kernel.<sup>14</sup>

## 2.4. Applications

Naturally, information implied in option prices about the state of the economy is interesting in and by itself – and surely for academics in particular. But a number of applications show that this information has wider relevance. First, national banks paid already attention to the risk-neutral distribution, which can be backed out from index option prices, see e.g. Bahra (1997)

---

<sup>14</sup> The empirical evidence is consistent with Branger, Hansis, and Schlag (2011); compare for the theoretical results also Chaudhury and Schroder (2015).

for the Bank of England. Under the assumption of a representative investor with a power utility, Sarantopoulou-Chioureaa and Skiadopoulos (2015) back out time-varying risk-aversion coefficients from data on S&P 500 index options. They then use the implied risk-aversion to predict real economic activity. Another recent, although rather superficial, study by Haas, Fajardo, and Rocha (2011) for the Banco Central do Brazil backs out the risk-aversion coefficient from Brazil real/US dollar exchange rates. Kelly, Pastor, and Veronesi (2015) analyze political uncertainty via its impact on equity option prices.

Second, applications in portfolio optimization utilize the forward-looking, risk-neutral distributions from option prices in order to construct superior portfolios. Ait-Sahalia and Brandt (2007) express a standard portfolio optimization problem in terms of the risk-neutral distribution instead of the more common dynamic programming approach. Under suitable assumptions, the two approaches are equivalent. Interestingly, the resulting optimal portfolios are quantitatively different from those based on standard assumptions about the return evolutions. DeMiguel, Plyakha, Uppal, and Vilkov (2013) use option implied volatilities and correlations for portfolio optimization. Using options implied information for portfolio optimization however restricts the set of possible stocks to the ones having liquid options.

Kostakis, Panigirtzoglou, and Skiadopoulos (2011) use the option-implied, risk-neutral distribution for the S&P 500 index and the assumption of an exponential or power utility to obtain forward-looking physical distributions. Their approach ignores the pricing kernel puzzle by design and also has a severe look-ahead bias as the risk aversion parameter is chosen so as to best fit the (supposedly unknown) future returns.<sup>15</sup> Their finding, that the forward-looking physical distributions then produces better portfolios than the historical

---

<sup>15</sup> The methodology is very similar to Bliss and Panigirtzoglou (2004). Zdorovenin and Pezier (2011) use a close variant, too, and are subject to the same critique as Kostakis, Panigirtzoglou, and Skiadopoulos (2011).

distribution, is thus not surprising. Unfortunately, they do not test for performance differences but only provide point estimates. For forecasting with option implied information we refer to the survey of Christoffersen, Jacobs, and Chang (2012).

## **2.5. The pricing kernel puzzle in exclusively forward-looking data**

The canonical papers on the pricing kernel puzzle mix backward-looking estimates of the physical distribution with forward-looking risk-neutral distributions in order to finally find the empirical pricing kernel as the ratio  $q/p$ . Inherently, any mismatch in time could contribute to the pricing kernel puzzle, and one has to worry if such mismatch spuriously created the pricing kernel puzzle. One way of using only forward-looking data is the approach of Bliss and Panigirtzoglou (2004) who assume a power utility function to find forward-looking physical distributions from estimated forward-looking risk-neutral distributions. Working out the quantiles of forward-looking, realized returns under the physical cumulative distribution function, they note that the collection of empirical quantiles should be standard uniformly distributed. The original formulation precludes the existence of a pricing kernel puzzle a priori by assuming a monotonically decreasing pricing kernel based on a power utility.

An obvious extension is to relax the assumption of a power or exponential utility and allow the pricing kernel more flexibility. Linn, Shive, and Shumway (2014) do exactly that by modeling the pricing kernel (assumed to be constant through time) through cubic B-splines. Starting with the risk-neutral distribution at an observation date, they divide by an assumed pricing kernel and obtain the physical distribution for that date. That allows them to work out the percentile of the realized market return over the next period given the physical cumulative distribution function. These percentiles across all observation dates should be uniformly distributed. Linn, Shive, and Shumway (2014) then optimize the parameters, which define the pricing kernel by fitting the moments of that uniform distribution via GMM. Surprisingly,

they find that the average pricing kernel for SPX options is monotonically decreasing while the average pricing kernel for FTSE options is u-shaped.

Cuesdeanu and Jackwerth (2016) revisit the problem of Linn, Shive, and Shumway (2014) but add a statistical test and can formally reject the null hypothesis of a monotonically decreasing pricing kernel. They thus contradict the findings of Linn, Shive, and Shumway (2014) and attribute this to (i) a lack of scaling so that the physical distributions of Linn, Shive, and Shumway (2014) are not integrating to one and (ii) a mismatch in their optimization (based on moments of the uniform distribution via GMM) and their measurement of fit (based on the Cramer van Mises statistic).

Sala and Barone-Adesi (2015) recognize the time mismatch, too, and propose to estimate, via a Bayesian procedure, a conditional physical distribution, which is based in part on a GJR-GARCH model using historical return data (based on the physical distribution) and in part on another GJR-GARCH model which has been fitted to option prices (based on the risk-neutral distribution). The latter risk-neutral distribution is given a new mean to account for the presence of a risk premium. The pricing kernel is finally obtained as the ratio of the risk-neutral and the conditional physical distributions and found to be monotonically decreasing. This is not too surprising given the circularity of basing the conditional physical distribution partially on (a mean-shifted version of) the risk-neutral distribution. Indeed, if the risk-neutral distribution were lognormally distributed, and all weight were given to the risk-neutral distribution in finding the conditional physical distribution, then the pricing kernel would be a ratio of two lognormals with different means; we would be back at a power utility and the associated decreasing pricing kernel.

### **3. Empirical Studies**

Most of the work on the pricing kernel puzzle investigates the S&P 500 index, and there are recent additions to this literature. More as an illustration than as a comprehensive study,

Jackwerth (2004) documents the pricing kernel puzzle for one single day in October 2003 for each of the markets for the DAX 30 (10/11), the FTSE 100 (10/14), and the Nikkei 225 (10/11), where the pricing kernel for the Nikkei turns out to be u-shaped.<sup>16</sup> A large number of studies have subsequently investigated if the pricing kernel puzzle also exists in other indices and have largely confirmed this finding for a number of large indices (e.g. the DAX and the FTSE). Little is known about the time-series properties of the pricing kernel puzzle. Finally, we turn to investigations of the pricing kernel puzzle in markets other than index markets. The main issue being here is that the pricing kernel is now the projection of the economy-wide pricing kernel onto the space of returns investigated (say foreign exchange). Depending on the correlation between the index (proxying for aggregate wealth) and foreign exchange (as the return under investigation), the pricing kernel puzzle might no longer show up in the data.

### **3.1. The S&P 500 index market**

A series of papers (the early draft of Shive (2003), then the extensions in Shive and Shumway (2004), and a shortened and tightened version in Shive and Shumway (2009)) investigate a large number of markets, including the S&P 500. So we will present the details only once and then only refer to the different markets later where appropriate. Shive (2003) uses a kernel regression of option prices directly (instead of the much better behaved implied volatilities, which only change within one order of magnitude across moneyness) and the Breeden and Litzenberger (1978) technique to obtain the risk-neutral distribution. For the physical distribution, she employs a kernel density estimator using five years of historical data. A third degree polynomial is fitted to the pricing kernel. She then investigates the local slope of the pricing kernel and employs a bootstrap test, which unfortunately is not detailed at all. The pricing kernel for the S&P 500 turns out to be locally increasing between February 1999 and November 2001. In Shive and Shumway (2004), the physical distribution is instead being

---

<sup>16</sup> See Section 3.6 for a discussion on tilde versus u-shaped pricing kernels.

obtained via a GARCH (1,1) and the pricing kernel is fitted to a power utility pricing kernel with the risk-aversion coefficient replaced by an affine function in the market return. The pricing kernel puzzle can be shown (via positive correlations between the underlying and the pricing kernel) in the extended version but not in case of power utility, when the risk aversion coefficient is constant and not a function of returns. The final draft of Shive and Shumway (2009) goes back to kernel density estimation for the physical distribution, although now with some volatility rescaling in order to obtain conditional estimates. The kernel regression of the option prices is replaced by the LOWESS smooth, a local polynomial regression where they unfortunately do not provide details on either polynomial order or bandwidth. The pricing kernel puzzle can be documented as the slope of the pricing kernel is locally increasing for the S&P 500.

In a more recent confirmation for one day of S&P 500 index options data at multiple maturities, Fengler and Hin (2015) work out the pricing kernel surface by fitting B-splines to option prices in order to find the risk-neutral distributions via Breeden and Litzenberger (1978) and fit a normal inverse Gaussian to obtain the physical distribution at multiple horizons.

Song and Xiu (2015) add information about the VIX level when estimating empirical pricing kernels for the S&P 500 using kernel based methods akin to Ait-Sahalia and Lo (2000). They confirm the pricing kernel puzzle unconditionally, but cannot establish it conditionally on high or low VIX levels. Thus, they speculate that stochastic volatility could be driving the pricing kernel puzzle but find that standard stochastic volatility option pricing models cannot generate the observed patterns.<sup>17</sup> See also Section 5.2.1 which suggests solutions to the pricing kernel puzzle based on volatility as a second state variable.

---

<sup>17</sup> A related observation by Boes, Drost, and Werker (2007) is that that the risk-neutral distribution, conditional on a low spot volatility, does not exhibit negative skewness.

### 3.2. The German DAX 30 index market

Concerning the German market, a number of papers from Humboldt University in Berlin have confirmed the pricing kernel puzzle for the DAX 30. Golubev, Haerdle, and Timofeev (2014) present a formal test of monotonicity of the pricing kernel (detailed below in Section 4.3) and detect the pricing kernel puzzle in the DAX in June 2002, but not in June 2000 or June 2004. Using the same data, Detlefsen, Haerdle, and Moro (2010) identify three different states of the economy: a bull market in March of 2000, a bear market in July of 2002, and a “sideways” market in June of 2004. They use the Heston (1993) model fitted to the options implied volatilities to obtain the risk-neutral distribution, which however loses some flexibility due to the rigid structure of the Heston (1993) model. For the physical distribution, they use a GARCH model, the Heston (1993) model, and the kernel density based on historical data. Using this setup, they confirm the pricing kernel puzzle for the bear and the sideways market, but not for the bullish where the pricing kernel is too stretched out to exhibit local non-monotonicity. Giacomini and Haerdle (2008) confirm the pricing kernel puzzle for the DAX in 2001 using a method similar to Ait-Sahalia and Lo (2000) where the risk-neutral distribution (at different horizons) is obtained by using Breeden and Litzenberger (1978) on options prices based on a smoothly estimated implied volatility surface. The physical distribution is fitted via a GARCH (1,1) process.

Shive (2003) and Shive and Shumway (2004, 2009) all find the pricing kernel puzzle to exist in the DAX data.

Much more interesting are the following two studies on the DAX which try to explain the time series properties of the pricing kernel puzzle. First, Giacomini, Handel, and Haerdle (2008) use tick data for the DAX from January 1999 to April 2002 and fit a GARCH model in order to obtain the physical distribution. The risk-neutral distribution estimation follows Ait-Sahalia and Lo (2000) by using a kernel regression of implied volatilities, followed by the conversion to option prices and an application of Breeden and Litzenberger (1978). Then,

time series of simple statistics of the pricing kernel plus the absolute and relative risk aversion functions at different maturities are being calculated and subjected to a principle component analysis. The principle components are finally regressed on returns on the DAX and on changes in at-the-money implied volatility. The main result seems to be the rather obvious finding that large changes in implied volatility lead to more volatile and time-varying pricing kernels.

Similarly, but using a slightly different technique, Grith, Haerdle, and Park (2013) use DAX data between April 2003 and June 2006. They fit a smoothing polynomial to the implied volatilities, translate those into option prices, and use Breeden and Litzenberger (1978) to obtain risk-neutral distributions. The physical distributions are based on two years' worth of historical returns via kernel density estimation. Finally, power utility functions are extended with four additional parameters (additive and multiplicative parameters inside and outside the power function) to allow for non-monotonic pricing kernels. Changes in these parameters and the location of the peak of the pricing kernel are being regressed on changes in the credit spread, the yield curve slope, the short interest rate, as well as the underlying return. Some correlations between those macro variables and the additional shape parameters are stated. The authors conclude that the locally risk loving behavior is pro-cyclical as the hump of the empirical pricing kernel seems to be more pronounced in calm periods.

These latter two studies on time series properties of the pricing kernel puzzle unfortunately do not relate their findings to clear economic arguments about the causes of time-variation in the pricing kernel puzzle. Much more insight relating to the best modeling of the pricing kernel puzzle could be gleaned from a more exhaustive search for explanatory factors of pricing kernel variation.

### 3.3. The British FTSE 100 index market

Liu, Shackleton, Taylor, and Xu (2009) were the first to perform a more extensive empirical pricing kernel study for the British market. Using FTSE 100 option data from 1993 to 2003 they fit a lognormal mixture distribution, a generalized beta distribution, and splines to the risk-neutral distribution. The physical distribution is estimated solely by a kernel density based on simulated paths of a GJR-GARCH(1,1) process whose parameters are estimated from the time series of the index. Their pricing kernels in their figure 2 exhibit the puzzling local increase, even though the authors downplay that fact. There are no statistical tests.

Consistent with this result, Shive (2003) finds the pricing kernel puzzle in the FTSE data when studying the subperiod 1999-2001, and Shive and Shumway (2004) find it for 1999-2003.

### 3.4. Other index markets

For other index markets, Shive and Shumway (2009) find evidence of the pricing kernel puzzle in the AMEX Japan index data. Shive and Shumway (2004) obtain the same result for the Swedish OMX index data.

Perignon and Villa (2002) estimate the relative risk aversion for the French CAC 40 index. They adopt the kernel method of Ait-Sahalia and Lo (1998) for the risk-neutral and physical distributions and derive the relative risk aversion function.<sup>18</sup> As the older options (prior to 1999) on the CAC 40 were American in style, they use only a short intraday sample from April 1999 to December 1999. The resulting relative risk aversion function is globally decreasing in wealth but locally increasing and negative in the wealthy states. Unfortunately, an empirical pricing kernel plot is not provided but the negative part of the risk aversion function would imply a u-shaped pricing kernel.

---

<sup>18</sup> The formula is very similar to the formula for absolute risk aversion in Equation (7). The only difference is that relative risk aversion has the return  $R$  as a multiplier.

Coutant (1999, 2000) uses earlier data on the CAC 40 from 1995-1996 to estimate risk aversion functions. To deal with American options, she applies the method of Melick and Thomas (1997), where the price of an American option is a linear combination of the upper and lower price bound of the option. The risk-neutral distribution is estimated by Hermite polynomials where the Gaussian distribution serves as a reference measure, see Madan and Milne (1994). The physical probability is estimated by quasi-maximum likelihood with the assumption that the process follows a geometric Brownian motion (Coutant 1999) or a GJR-GARCH process with t-distributed errors (Coutant 2000). The resulting absolute risk aversion functions in Coutant (1999) are u-shaped. In the 2000 version however, she restricts the representative investor to have power preferences and can thus only show decreasing pricing kernels.

Lioui and Malka (2004) work with Israeli data on the TA-25 index. Due to their restrictive power utility assumption, they cannot document the pricing kernel puzzle.

Haerdle, Grith, and Mihoci (2014) detect the pricing kernel puzzle in the cross section of equities for Australia, Germany, Japan, Switzerland, the United Kingdom, and the United States. Using the market model of Grith, Haerdle, and Kraetschmer (2016), where the representative investor exhibits time separable, state dependent utility, the pricing kernel is parameterized by a downward sloping segment, which then jumps up or down, followed by another downward sloping segment for returns higher than some breakpoint  $x$ . This pricing kernel is estimated via GMM by means of an Euler equation and the optimal switching point is determined by a grid search. Considering the 20 largest stocks for each country, they find the evidence of the pricing kernel puzzle as the estimated pricing kernels jump up around the optimal switching point.

### 3.5. Non-index asset markets

First a word of caution on computing the empirical pricing kernel for non-index assets altogether. If one adheres to some notion of preferences over consumption, then a concentration on the index makes much sense. After all, consumption should be correlated with wealth and that in turn is driven to a large extent by the evolution of large indices such as the S&P 500. But considering some mildly correlated asset (say gold which is typically not much correlated with the stock market), one could compute a pricing kernel for the gold market from historical gold returns and options on (futures on) gold. But for the representative investor, we have no clear prediction of the shape of such pricing kernel in the gold dimension: a low gold price is not related to low stock market prices (poor state of the world, low consumption, high risk aversion) nor is the opposite true for high gold prices. Thus, pricing kernels on non-index assets might well turn out to be disappointingly flat and with little room for interpretation. The situation would be different for asset classes more highly correlated with the index. Moreover, for a careful, bivariate analysis of the pricing kernel puzzle, one would need to estimate bivariate risk-neutral distributions, which is exceedingly difficult as there are few options written on both assets at the same time (knowing only options on one asset and options on the other asset separately is typically not enough), and bivariate physical distributions. Jackwerth and Vilkov (2015) have recently made inroads here in estimating the bivariate risk-neutral distribution on the S&P 500 and the VIX, using longer-dated options to circumvent the above problem in this special set-up.

Considering non-index asset classes, the individual stocks take up a halfway position as they are the constituents of the index. Ni (2009) and Chaudhuri and Schroder (2015) analyzed individual stock options within the S&P 500. Chaudhuri and Schroder (2015) find evidence of return patterns compatible with the pricing kernel puzzle and criticize the earlier paper of Ni (2009), which cannot find such evidence, for methodological reasons. Details can be found in Section 2.3. Similarly, the work of Figlewski and Malik (2014) is based on option data on

exchange traded funds having the S&P 500 as an underlying. Due to the high correlation with the S&P 500, we do not really view this exchange traded fund as a non-index asset. Not surprisingly, their work finds non-monotonic pricing kernels. By considering exchange traded funds that aim to provide (i) twice the return on a long position in the S&P 500 and (ii) twice the return on a short position, they also contribute to the literature on heterogeneous investors and the pricing kernel puzzle, see Section 5.1.1.

Turning to commodities, Shive and Shumway (2009) find a u-shaped pricing kernel for futures options on corn, live cattle, lean hogs, soybean, and wheat. The authors interpret these results as evidence of biased beliefs as their expectation was that pricing kernels, which are estimated in dimensions uncorrelated with the index (e.g. agricultural products), should be essentially flat. But since agricultural products have values to the farmers, who might not be perfectly diversified in their investments, one could alternatively expect some slightly sloping pricing kernel, too.

Concerning foreign exchange markets, Haas, Fajardo, and Rocha (2011) use the parametric transformations from Liu, Shackleton, Taylor, and Xu (2007) to back out the risk-neutral and physical distribution of the Brazilian Real/US dollar exchange rate. Using data from 1999 to 2011, they estimate a relative risk aversion coefficient of about 2.7. For July 2006, they show in their graph 1 two probability distributions, which exhibit the pricing kernel puzzle.

Investigating interest rates, Li and Zhao (2009) estimate the risk-neutral distribution implied in interest rate caps (a cap is a portfolio of call options on an interest rate) by extending the locally polynomial estimator of Ait-Sahalia and Duarte (2003). While the latter regress call prices solely on moneyness, Li and Zhao (2009) allow prices also to depend on the slope and volatility of the LIBOR rates. The physical distribution is estimated by the historical LIBOR rate with the kernel method of Ait-Sahalia and Lo (2000). The resulting pricing kernel exhibits a u-shape where the left part of the kernel reflects investors' fear of a recession and the right part fear of inflation. The authors note that the u-shape is more pronounced at longer

maturities, since the physical distribution is more compact for 4 and 5 years than for 2 and 3 years due to the mean reversion of interest rates. Liu, Kuo and Coakley (2015) also investigate interest rates. Using options on futures on the 6-month LIBOR, they estimate a pricing kernel, which is based on a power utility but comes with a three parameter, flexible adjustment function. Estimating the pricing kernel via GMM leads to the pricing kernel puzzle, whereas estimating it based on the second Hansen-Jagannathan distance leads to monotonically decreasing pricing kernels.

Regarding inflation, Kitsul and Wright (2013) estimate an empirical pricing kernel on inflation by using caps and floors on inflation. Although this market is not as liquid as the related TIPS and inflation swap market, it is growing fast and mutual funds and insurance companies invest into it, as the authors argue. Prices are available from 2009 to 2012. Butterfly spreads determine the prices of the Arrow-Debreu securities (i.e. state prices), which in turn can be translated into the risk-neutral distribution through normalization. After dividing the risk-neutral distribution by the historical inflation distribution, the resulting pricing kernel is u-shaped, and the authors provide some robustness tests. While it is intuitively clear that states of high inflation/deflation are disliked (and thus command high levels of marginal utility), it is not apparent how the inflation states are connected to aggregate wealth or the market index.

For volatility, Song and Xiu (2015) and Bakshi, Madan, and Panayotov (2014) find u-shaped pricing kernels using options on VIX.

Covering many different asset classes, Chernov (2003) calibrates the S&P 500, two further equities, gold futures, and T-bills to a highly parameterized multivariate extension of the Heston (1993) model to capture many market segments. A more general economy-wide pricing kernel should result, in contrast to other studies, where often only the projection of the pricing kernel onto the index is estimated. However, the estimated pricing kernel turns out to be a highly non-linear function of the returns on the S&P 500, suggesting the presence of the

pricing kernel puzzle in the data. As a possible explanation, Chernov (2003) suggests habit formation; the later model of Chabi-Yo, Garcia, Renault (2008) picks up on this issue again, see for details in Section 5.2.1.

The key point concerning non-index asset classes is that pricing kernels can be estimated by “brute force” for other asset classes at all. Less clear is, what such pricing kernels teach us about the main dimension of interest, aggregate consumption, unless there is detailed information about the correlation between the asset returns and consumption growth (often proxied by returns on the index).

### **3.6 The shape of the empirical pricing kernel**

So far, no study examined the reasons for some researchers finding the empirical pricing kernel to be tilde-shaped while others finding a u-shaped pattern. Are u-shaped pricing kernels simply right shifted versions of the tilde-shaped ones? In that case, the right-most part of the tilde shape is simply clipped off, leaving a u-shape. Or, is the noise in the estimation of the far right tails of the risk-neutral and physical distributions responsible for the u-shaped pricing kernels being non-decreasing at the high moneyiness levels?

In order to shed some light on this issue, we perform a small empirical study by estimating a time-series of non-overlapping monthly S&P 500 pricing kernels from January 1987 to August 2015. The physical distributions are obtained by calibrating a GARCH(1,1) on a monthly basis using the past 3.500 daily returns (some 14 years of data). The risk-neutral distributions are obtained by applying the fast and stable method of Jackwerth (2004). The equity risk-premium is set to 5%. Each month we track the moneyiness level at which the local maximum in the empirical pricing kernels occurs. Whenever the local maximum is at a moneyiness level above 1.10 or whenever the pricing kernel is globally u-shaped, we assign a value of 1.10.

[Figure 4 about here]

Figure 4 shows the resulting time-series of moneyness levels for the local maxima. Looking at the graph, one would expect that this time-series is mainly driven by a latent volatility factor, and, indeed, the VIX has a correlation of 0.57 with the location of the local maximum. Hence, studies using mainly calm periods should be more likely to find tilde-shaped pricing kernels than u-shaped ones, even though both shapes are driven by the same puzzle. Our findings from this simple exercise are in line with studies using volatility as a second state variable to explain the pricing kernel puzzle, see Section 5.2.1. Nevertheless, we do not claim that volatility is the ultimate reason for finding different pricing kernel shapes but rather leave this task to future research. Further minor factors that add uncertainty to the pricing kernel estimation in terms of the location of the local maximum are potentially:

1. A misestimation of the equity premium: Some authors fix the equity premium, which is the difference between the means of the physical and risk-neutral distributions, to be constant over their sample. Others estimate the equity premium with a rolling window.
2. An over-/underestimation of the right tail of the physical distribution due to rarely observed, extremely positive returns.
3. An over-/underestimation of the right tail of the risk-neutral distribution due to the fact that deep out-of-the money calls are not as heavily traded as their out-of-the money puts counterparts, which are used for the estimation of the left tail of the risk-neutral distribution.

## **4. Econometric assessment of the pricing kernel puzzle**

Is the pricing kernel puzzle a statistically established fact or simply a spurious finding?

### **4.1. Sample variation bounds**

The canonical models provide bounds around the pricing kernel estimates simply based on the sample variation of the inputs, namely, historical returns and option prices, and those bounds

do not constitute formal tests of monotonicity. Using the bounds suggests that the estimated pricing kernels exhibit local increases exceeding those bounds. The main finding of Jackwerth (2000) is presented in his figure 3 where the risk aversion functions are negative by more than two standard deviations for the sample March 19, 1991-December 29, 1995. Moreover, Jackwerth (2000) reports a number of robustness tests which do not change the finding significantly, namely, different lengths for the historical sample, elimination of low-strike option prices, changes to the equity risk premium, changes to the kernel bandwidth, and changes to the time-to-maturity. Assessing the statistical significance of the pricing kernel puzzle in Ait-Sahalia and Lo (2000) is somewhat complicated as it involves a statement about the propensity of the supposedly downwards sloping pricing kernel to have locally increasing segments. The empirical pricing kernel of Ait-Sahalia and Lo (2000) does indeed increase in their figure 3 as the S&P 500 increases from 400 to 435. They also provide the 5% and 95% quantiles around their pricing kernel, and, by visual inspection, the upper quantile at an index value of 400 is very close to the lower quantile at an index value of 435. While this argument is not a formal statistical test, it is still highly suggestive of the presence of the pricing kernel puzzle in Ait-Sahalia and Lo (2000). The third canonical paper, Rosenberg and Engle (2002), finds the pricing kernel puzzle in their figure 6, which shows a clear local increase in the pricing kernel beyond the two standard deviation bounds around the empirical pricing kernel for the years 1991 to 1995 in the S&P 500 market.

#### **4.2. Perturbations of options data**

Going one step further, Bliss and Panigirtzoglou (2002) investigate the impact of a uniformly distributed error in the option prices within the bid-ask spread on the risk-neutral distribution. Haerdle, Okhrin, and Wang (2015) take up this concept and investigate the impact of errors in option prices or implied volatilities on the empirical pricing kernel. Both the risk-neutral distribution and the physical distribution are being obtained through kernel based techniques,

which allow the authors to describe the uniform confidence bands around the empirical pricing kernel in statistical terms.

A careful study of small sample noise in both the physical and the risk-neutral distribution is Leisen (2014).<sup>19</sup> He finds that spurious non-monotonicities can arise for simulations of power utility pricing kernels. The problem is particularly relevant if the physical distribution is based on historical samples of only 48 monthly returns, and the situation improves much once a GARCH(1,1) model is estimated. Also, the risk-neutral distributions are based on Ait-Sahalia and Lo's (2000) kernel-based method, which is noisier than other methods for backing out risk-neutral distributions from option prices.

### 4.3. Statistical tests

The complicated issue of formally testing for locally increasing segments of the estimated pricing kernel has been taken up in Golubev, Haerdle, and Timofeev (2014) under the strong assumption of iid realized returns. The idea is to map the problem to an exponential model and check for pricing kernel monotonicity between any two realized returns in the sample. The fairly complicated test then considers the joint distribution of monotonicity violations across all possible combinations of observed returns. Applying their test to the DAX index during the summers of 2000, 2002, and 2004, monotonicity could be rejected at the 10% significance level in 2002, but not for the years 2000 and 2004.

Another test is Härdle, Grith, and Mihoci (2014), which uses the market model of Grith, Haerdle, and Krätschmer (2016). Here, the pricing kernel is parameterized as two decreasing segments with some breakpoint where the pricing kernel jumps up or down. Comparing GMM estimates of the restricted model (the two segments join smoothly in a decreasing

---

<sup>19</sup> See also Lioui and Malka (2004) for reported differences due to using either only call or only put options.

manner) versus the unrestricted model, the authors employ a so-called D-test and reject pricing kernel monotonicity in typically four out of five cases.

A further attempt at designing a formal statistical test is Beare and Schmidt (2015) who base their test on the equivalence of the monotonicity of the pricing kernel and the concavity of the ordinal dominance function. The latter function is the cumulative risk-neutral distribution of the quantile function of the physical distribution. They find that in about half the months from 1997-2009, the pricing kernel puzzle can be detected at the 5% significance level.<sup>20</sup>

Cuesdeanu and Jackwerth (2016) suggest a simpler test based on risk-neutral distributions, which have been divided by some pricing kernel to find the physical probability distributions at each observation date. Working out the quantiles of the observed market returns while under the physical cumulative distribution function, the quantiles throughout the sample should be standard uniformly distributed. The authors then optimize several test statistics of uniformity while either restricting or not restricting the pricing kernel to be monotonically decreasing. The discrepancy in optimized test statistics can then be tested against its simulated distribution. Cuesdeanu and Jackwerth (2016) confirm the presence of the pricing kernel puzzle in the S&P 500 index options data from 1987 to 2015. Note that the paper, as opposed to the earlier canonical studies, uses only forward-looking data, namely, the physical returns are forward-looking and no longer based on historical samples.

Another attempt to statistically test for locally increasing empirical pricing kernel is Shive (2003) who obtains empirical pricing kernels from essentially the method of Ait-Sahalia and Lo (2000) while limiting the options to one monthly cross section at the time whereas Ait-

---

<sup>20</sup> Leading even further afield is the approach of Bakshi and Panayotov (2008) to investigate option mispricing via the law of one price. Here, the monotonicity of the pricing kernel is no longer required and the pricing kernel can even go negative. Thus, the pricing kernel puzzle could not even be detected, even if it existed in the data.

Sahalia and Lo (2000) use all option prices during the whole year of 1993. Shive (2003) then fits the obtained pricing kernels to a third order polynomial, takes the first derivative, and checks at varying index levels for a positive slope of the pricing kernel. The statistical significance is being argued through a bootstrapped test where unfortunately it is completely unclear what and how is being bootstrapped. Shive and Shumway (2004) closely follow the methodology of Shive (2003) except that the empirical pricing kernel is now in one version used as is and in a second version also fitted to an eighth order polynomial. Local increases in the pricing kernel are being established through local correlation coefficients with the index on specific intervals. Again, the test uses a bootstrap approach, which has not been sufficiently detailed.

Given that many monotonicity tests are problematic or at least rather technical, it might be interesting to base a test on the formula for absolute risk aversion in equation (7) which we repeat here for convenience.

$$ARA_i = -\frac{U''(R_i)}{U'(R_i)} = \frac{p'}{p} - \frac{q'}{q} \quad (7)$$

The advantage would be that it might be simpler to test for negativity of absolute risk aversion rather than non-monotonicity of the pricing kernel.

Note also the statistical test of Chaudhury and Schroder (2015), which is an indirect test as it utilizes returns on options instead of the usual technique based on index returns and option prices. Details can be found in Section 2.3. Their test confirms the presence of the pricing kernel puzzle for the S&P 500 index.

Patton and Timmermann (2010) suggest a monotonic relation test for asset returns. They claim that their test can be easily implemented for validating pricing kernel monotonicity. To apply their test one would presumably discretize the empirical pricing kernels at 9 moneyness levels and assume monotonicity under the null hypothesis. It would be interesting to see the implications of their test when taking it to the data.

## 5. Solutions

Considering the empirical evidence and the statistical tests so far, it emerges that the pricing kernel puzzle seems to be present in the data. Taking it to be literally true, however, seems like a naïve interpretation. In that case, the representative investor of the simple economy in Section 1 would need to have a convex segment in the utility function, akin to the Friedman and Savage (1948) utility function in Figure 5.<sup>21</sup> For a representative investor, this is hard to reconcile with equilibrium. It would mean that the representative investor was better off by not investing into states of world, where the index pays off when the utility function is convex. Rather, the representative investor would prefer a lottery over the two adjacent states (0.9 and 1.1 on the return axis of Figure 5) where the utility function turns concave again. But such avoidance of states jars with the notion that the representative investor needs to hold all assets by definition. Rather, security prices need to adjust so that the representative investor is willing to hold all assets in equilibrium. This point is made more rigorously in Hens and Reichlin (2013).

[Figure 5 about here]

Looking at the equilibrium problem for a different angle, Beare (2011) works out, based on some earlier results by Dybvig (1988), measure preserving derivatives which any investor

---

<sup>21</sup> Note that Friedman and Savage (1948) introduced their utility function for individuals and not for the representative investor. In particular, their concern was with small stakes gambling such as buying a lottery ticket. Chetty and Szeidl (2007) provide a microeconomic motivation for Friedman-Savage utility via consumption commitments (e.g. housing), for which the spending cannot easily be adjusted. Again, this is a model of individual investors and it is not obvious that the convexities would survive aggregation to a representative investor. See Ingersoll (2017) for related results on another partially convex utility function, namely cumulative prospect theory.

should prefer to investing into the market (see also Rieger (2011)). Their prices are less than the price of the market in times where the pricing kernel puzzle exists, and Beare and Schmidt (2014) show that returns on an option portfolio exploiting this circumstance actually stochastically dominate market returns. While such disequilibrium could well exist for some period in time, it is hard to see how such situation could persist unabated ever since the crash of 1987. As a result, researchers turned their interest away from disequilibrium and rather asked which richer economic settings could give rise to the pricing kernel puzzle. We will now investigate richer models using only a single state variable, then the important class of models with more than one state variable, before turning to behavioral and sentiment models, and finally to ambiguity aversion models.

## **5.1. Models with a single state variable**

A first step at resolving the pricing kernel puzzle was made in 2004 by Brown and Jackwerth (2012). Their idea was to allow the volatility of the index to be a function of a weighted average of historical index values. This small step away from a fixed volatility model leads to a pricing kernel exhibiting the puzzling non-monotonicity, although, when using realistic parameters, just to a very small degree.

### **5.1.1. Heterogeneous investor models**

Another way of extending the simple setting of Section 1 is to replace the representative investor with several (classes of) heterogeneous investors. Bakshi and Madan (2007) assume heterogeneity in beliefs in a complete market. Investors have different subjective distributions (instead of homogeneous belief in the physical distribution); consequently, investors expecting positive returns are long in the market, while investors expecting negative returns are short. The aggregation of both groups of investors leads to a u-shaped pricing kernel. Ziegler (2002) uses a very similar set-up and can show that the risk-neutral distribution can

become even bi-modal, if the beliefs are strongly heterogeneous.<sup>22</sup> He documents negative relative risk aversion functions, consistent with the pricing kernel puzzle.

The ideas of Bakshi and Madan (2007) are further developed in Bakshi, Madan, and Panayotov (2010), where 3 groups of investors have power utility with the same risk-aversion parameter ( $\gamma = 12$ ) and there exists within each cohort a fraction of investors shorting the market (i.e. six different groups of investors). The resulting pricing kernel is u-shaped by construction. The proportion of pessimists, shorting the market, needed to make the model work, seems to be quite large (10%-30% across groups) and an explicit calibration to prices is missing.

Ziegler (2007) examines a complete market with multiple investors (as opposed to his earlier work, Ziegler (2002), with just two extreme investors) and assumes that the index is a good proxy for consumption. His results indicate that neither (i) aggregation of (heterogeneous) preferences, (ii) misestimation of beliefs, nor (iii) heterogeneous beliefs can lead to reasonable explanations of the pricing kernel puzzle. He shows that, given reasonable individual utility functions, aggregation of heterogeneous preferences alone cannot explain the puzzle as the economy-wide risk-aversion inherits the behavior of the individual risk-aversions. In order to deal with misestimated beliefs, the stochastic volatility, stochastic jump model of Pan (2002) is considered.<sup>23</sup> Fitting the model to the data and assuming that investors

---

<sup>22</sup> In such setting, Shefrin (2008a and 2008b) coins the term sentiment for the ratio of the mixture of the different subjective distributions and the physical distribution. His ideas become clearer when one assumes that the shapes of the subjective distributions and the physical distributions remain the same but the means is low for the pessimists, high for the optimists, and in between for the physical distribution, see Shefrin (2008b, figure 1).

<sup>23</sup> Although the model captures stochastic volatility and jumps, the risk-aversion functions turn negative for high return states. Such behavior contradicts the standard assumption of a

have homogeneous beliefs but cannot estimate them correctly, Ziegler (2007) argues that the resulting misestimation is too severe to be credible.

When allowing for heterogeneity among beliefs, Ziegler (2007) needs a large share of investors with very pessimistic beliefs to explain the puzzle, similar to the situation in Ziegler (2002). Hence, a fat left tail can only be captured if some investors expect extremely negative returns. However, a setting with three groups of investors is only capable of generating the pricing kernel puzzle if two of the groups are unrealistically pessimistic. Ziegler (2007) then already suggests that a solution of the pricing kernel puzzle needs to go beyond the rather simple setting of a complete, frictionless market with a single state variable.

Siddiqi and Quiggin (2016) model an economy where agents display differential awareness, which occurs when an investor changes his set of possible outcomes. They contradict Ziegler (2007) by claiming that already a small distortion in beliefs can cause the pricing kernel puzzle, which, for example, obtains under the following conditions: stock and option traders might completely ignore states of the world, where the stock price would fall below some threshold. After receiving (pessimistic) news, the stock trader is aware of a large decrease in the stock price, while the option trader ignores this fact. While being at times possible for the less traded single stock options, it is hard to believe that differential awareness is plausible for the index over extended periods of time.

In a two dates exchange economy with a finite number of states, Hens and Reichlin (2013) systematically examine violations of three basic assumptions of their model (namely, risk-averse behavior, unbiased beliefs, and complete markets). All three relaxations can then generate the pricing kernel puzzle. Quite obviously, allowing for a partially convex utility

---

risk-averse representative investor, which leads to the question, if stochastic volatility, stochastic jump models are typically incapable to fit the historical risk-neutral and physical distribution simultaneously.

function (e.g., Friedman and Savage (1948)) will generate the pricing kernel puzzle by design. However, a representative investor would not allocate wealth to states where the utility function is convex, and the relaxation is thus unrealistic.

Biased beliefs are modeled in two ways by Hens and Reichlin (2013). First, as humans tend to overweight less probable extreme events, beliefs could be systematically distorted according to the model of Tversky and Kahneman (1992). Second, beliefs could be biased as different investors fashion different subjective forward-looking distributions based on the same historical return distribution. In isolation, both types of biased beliefs are incapable to explain the puzzle. However, by combining both types, the authors can generate the pricing kernel puzzle, although only at the cost of assuming a negative expected mean return for the representative investor. Finally, Hens and Reichlin (2013) introduce background risk as a form of market incompleteness. In a simple four state example, two investors facing background risk individually can generate the pricing kernel puzzle.

The plausibility of heterogeneous beliefs and preferences is considered in Figlewski and Malik (2014) from an empirical point of view. The authors examine options on an exchange traded fund replicating the S&P 500 (SPY), on one that aims to provide the return on a two-times long position in the index (SSO), and on one that aims to provide the return on a two-times short position (SDS). Presumably, optimistic investors will buy the SSO fund; pessimistic investors the SDS. The paper then studies two extreme cases: (i) pricing kernels could be the same but not subjective distributions or (ii) pricing kernels could differ but all investors share the belief in the same physical distribution. It turns out that setting (i) explains the data better. Unfortunately, the set-up does not allow for intermediate settings between the extreme cases. Last, it is suggested that preferences within each group should be constant over time and the daily change in expectations stems from a change in the risk-neutral distributions.

### 5.2.2. Piece-wise approximations of the pricing kernel

As opposed to many of the above papers, which use equilibrium approaches to aggregate the individual investors' utility functions to a market-wide pricing kernel, a surprising number of papers use rather ad-hoc assumptions in order to aggregate utility functions. Detlefsen, Haerdle, and Moro (2010) suggest a harmonic mean of individual utility functions, without referring to a market equilibrium. While such approach can be used in order to allow for flexible pricing kernels, which exhibit increasing sections, one should not take the model literally. Haerdle, Kraetschmer, and Moro (2009) was subsumed by Grith, Haerdle, and Kraetschmer (2015), who piece together the pricing kernel from many segments which (between referent points) are decreasing but can jump upwards at the reference points. Investors are allowed to have different reference points. Given a sufficient number of such reference points, the authors can generate a flexible pricing kernel specification, which can exhibit increasing parts. One can study its piece-wise nature in their figures 2 and 3 in detail. In their empirical section they find that the local maximum of the pricing kernel near at-the-money is more pronounced when the variance risk premium is low.<sup>24</sup>

In conclusion, it seems rather hard to explain the pricing kernel puzzle with only one state variable. Moreover, there is always the nagging doubt of how a locally increasing segment of the pricing kernel can be reconciled with equilibrium. A representative investor would not want to hold securities that pay off in such states and models with several (groups of) investors need to have rather strongly diverging beliefs (very pessimistic investors vs. rather optimistic ones), while ignoring the large mass of moderate investors in the middle.

---

<sup>24</sup> Here it would be worthwhile to have a model that produces a pricing kernel, which is a function of the variance risk premium, allowing the degree of non-monotonicity to change with the variance risk premium. See Section 5.2.1 for such models with volatility being a second state variable.

## 5.2. Models with several state variables

One way out is being hinted at by Brown and Jackwerth (2012) who introduced the (weighted) average historical volatility as a new variable. While it is still deterministically driven by the return process (which technically makes it a single state variable model), it opens up the perspective of introducing additional state variables. The pricing kernel would then exist across those several dimensions, and the pricing kernel projected onto the return dimension might then exhibit the pricing kernel puzzle.<sup>25</sup>

### 5.2.1. Models with volatility as a second state variable

Christoffersen, Heston, and Jacobs (2013) stay close to the above idea and extend the Heston and Nandi (2000) model by introducing a variance risk premium in addition to the equity risk premium. The pricing kernel is now a function of returns and volatility. When projected onto returns only, by construction, a u-shaped pricing kernel emerges whenever the variance premium is negative. Fitting this GARCH model to the historical time series and a cross section of Wednesday options on the S&P 500 from 1996 to 2009 while allowing for a variance premium, and hence for a u-shaped pricing kernel, improves the risk-neutral and physical fit substantially. The quadratic functional form of the pricing kernel is rigidly assumed by the model and at times does not fit the empirical tilde-shaped pricing kernel in the empirical section of their paper.

In a follow-up study, Babaoglu, Christoffersen, Heston, and Jacobs (2016) set up a nested model and show that u-shaped pricing kernels are more valuable for fitting the option and return data than fat-tails and a second volatility component. In a related setting, Bollerslev and

---

<sup>25</sup> A number of papers show that such additional state variables seem to be empirically needed in order to explain option prices, see for example Buraschi and Jackwerth (2001), Coval and Shumway (2001), and Constantinides, Jackwerth, and Savov (2011).

Todorov (2011) show how investor fear contributes simultaneously to the equity and variance risk premia. They also find evidence for a u-shaped pricing kernel.

Chabi-Yo (2012) shows that a recursive small-noise expansion results in a pricing kernel that incorporates stochastic volatility, stochastic skewness, and stochastic kurtosis, while an ordinary Taylor expansion would lead to a pricing kernel, which is a polynomial in the market return.<sup>26</sup> Using French's 30 monthly industry portfolios, he recovers the higher moment preferences of the representative investor. His empirical pricing kernel is a function of volatility and return. Holding volatility fixed, it is monotonically decreasing in the market return. Yet, when projected onto the market return only, the empirical pricing kernel shows the puzzling behavior. For robustness, he shows that the pricing kernel projected onto the market return exhibits a similar shape if it is estimated with the S&P 500 option data rather than industry portfolio returns.

While most of the literature on heterogeneous beliefs and the pricing kernel focused on disagreement on the expected return (see e.g. Ziegler (2007) and Hens and Reichlin (2013)), Bakshi, Madan, and Panayotov (2014) consider heterogeneity with respect to future volatility and allow the investors with exponential utility to have different levels of risk-aversion, too. As a result, they obtain a u-shaped pricing kernel in the volatility dimension from options on VIX.<sup>27</sup> In contrast, most standard models imply that the pricing kernel is monotonically

---

<sup>26</sup> A group of papers, such as Dittmar (2002), assumes the pricing kernel to be some Taylor series expansion of returns. Here it is not quite clear if such, largely parametric, assumption of a functional form constitutes an economic model already or is merely a numerical fitting technique, such as in e.g. Rosenberg and Engle (2002).

<sup>27</sup> Song and Xiu (2013) confirm u-shaped volatility pricing kernels and find mostly monotonically decreasing pricing kernels in returns, when conditioning on volatility. They speculate that VIX might thus be a potentially important second state variable but do not

increasing in volatility. Therefore, the model could potentially solve the pricing kernel puzzle as returns around zero are associated with low volatility, and low volatility on the other hand is associated with an increasing pricing kernel. Unfortunately, the paper does not explore this intriguing aspect.

Kiesel and Rahe (2015) extend the stochastic volatility model of Heston (1993) by modeling the risk-neutral variance process with a long-term volatility target, which is yet again driven by another Ornstein-Uhlenbeck process. The physical process they leave unchanged. As a result, the implied pricing kernel can exhibit increasing sections.

### 5.2.2. Regime-switching models

Garcia, Luger, and Renault (2003) first introduced regime switches in the fundamental state variables of an equilibrium model and used this model to price options. Extending this work, Chabi-Yo, Garcia, and Renault (2008) show that the pricing kernel puzzle can be explained by regime-switches in some latent state variable, which in turn drives fundamentals (the joint distribution of the pricing kernel and returns), preferences, or beliefs. Their model uses a recursive Epstein-Zin (1989) utility for modeling the first two aspects (fundamentals and preferences), and an external habit model, which is based on Veronesi (2004) and Campbell and Cochrane (1999), for modeling the third aspect (beliefs).<sup>28</sup> The intuition is that,

---

develop a model, which would incorporate this insight. However, since their risk-neutral distributions are backed out from different snapshots in time, non-decreasing pricing kernels could average out so that they look as if they are monotonically decreasing, see Beare and Schmidt (2014).

<sup>28</sup> A further, purely theoretical, model using additive habit formation is Han and Turvey (2010). Such assumption has the implication that consumption needs to increase over time, which is hard to reconcile with reality. Investor risk aversion can thus vary through time and

conditional on the latent state variable, the pricing kernel is not violating the standard monotonicity assumption, whereas a projection of the pricing kernel onto returns leads to a locally increasing pricing kernel. Indeed, a simulation with hypothetical parameters can reproduce the desired shapes for the conditional and unconditional pricing kernels. One can note in the figures that the modeled pricing kernels often do not match the empirical pricing kernels in shape and magnitude. A more full-fledged empirical exercise might be able to improve the fit. A related model with uncertainty about endowment growth is Lundtofte (2010), where the investor does not know which of several growth rates (with associated probabilities) is the true one, before learning about the final value after some (short) time. Assuming reasonable parameters and two growth rate states, the model is able to generate a locally increasing pricing kernel.

Benzoni, Collin-Dufresne, and Goldstein (2011) extend the long run risk model of Bansal and Yaron (2004) where investors have Epstein-Zin preferences by adding learning and jumps to the expected consumption growth and consumption volatility processes. The jumps are again driven by a regime-switching process as in Chabi-Yo, Garcia, and Renault (2008); a modeling choice which allows consumption itself to stay relatively smooth. However, the learning aspect is the most exciting contribution of the model.

It has been known for a long time that the pricing kernel puzzle only emerged after the crash of 1987 and that it did not exist in the options data before the crash, see Jackwerth (2000). The important novelty of Benzoni, Collin-Dufresne, and Goldstein (2011) is then that investors can update their beliefs due to the crash to exhibit very persistent “crash-o-phobia” thereafter, even though the crash did not repeat itself for quite some time. As a result, the

---

even become negative. The paper argues that a locally increasing pricing kernel can result. As the article is purely theoretical, it remains unclear if this model can reproduce the empirical findings with reasonable parameters.

implied volatility smile changes from almost flat to steeply skewed for the index<sup>29</sup> while aggregate consumption does not change by much. The model is calibrated to a number of stylized facts about the S&P 500 data. Unfortunately, it does not show the model pricing kernel in the return dimension, and one cannot easily determine if it exhibits the pricing kernel puzzle; the pricing kernel in the dimension of consumption is monotonically decreasing by assumption.

### **5.3. Behavioral and sentiment models**

After first looking at demand based models, we next turn to models with probability weighting.

#### **5.3.1. Demand based models**

Bollen and Whaley (2004) come tantalizingly close to tackling the pricing kernel puzzle in their study of demand for out-of-the-money put options. They first establish that the physical distributions for individual stocks and for the S&P 500 index are not that different. They then turn to the implied volatility smiles, which are mildly u-shaped for individual stock options and steeply skewed for the index. Their explanation is that strong investor demand for portfolio insurance exists for out-of-the-money index puts, but is weaker for individual stock

---

<sup>29</sup> Furthermore, their model is able to explain the flatter implied volatility smiles for individual options, see Bakshi, Kapadia, and Madan (2003). Another approach of reconciling the steep index smiles with the flat stock option smiles is Branger and Schlag (2004) who introduce jumps and the associated premia. Unfortunately, they do not take the model to the data in order to see if the premia are realistic, nor do they investigate the pricing kernel puzzle. Chaudhury and Schroder (2015) go further in this respect in a rather similar setting, stressing that jumps in the individual stocks need to occur simultaneously. Then, the individual stocks do not exhibit the pricing kernel puzzle while the index does.

option puts. The high demand for out-of-the-money index puts by institutional investors is only met with supply by the market makers at rather high prices, moving the implied volatilities up, and causing the steep smile.<sup>30</sup> Having thus explained the cause of the steep index smile, they unfortunately do not connect their story to the pricing kernel puzzle, even though just one final argument is required. Namely, as the steep index smile leads to a left-skewed, leptokurtic risk-neutral distribution, the pricing kernel puzzle emerges once the risk-neutral distribution is being divided by the more normally distributed physical distribution. For the individual stock options, the mild smile leads to rather normally distributed risk-neutral distribution in the dimension of individual stock returns and, thus, the pricing kernel puzzle does not emerge when dividing by the physical distribution.

Motivated by these empirical results, Garleanu, Pedersen, and Poteshman (2009) develop a demand based option pricing model by departing from no-arbitrage principles, considering the options market as being separated from the underlying, and highlighting the importance of the market maker. In the presence of jumps and stochastic volatility, market makers cannot fully hedge their exposures and will demand higher prices for options paying off in states where hedges are critical. Hence, the resulting implied volatility smile is increasing in regions where hedging is more difficult for the market maker, which mainly concerns out-of-the-money puts. Similarly to Bollen and Whaley (2004), they find that option end-users are typically long index puts and short single stock calls. Again, an explicit treatment of the pricing kernel is missing.

Hodges, Tompkins, and Ziemba (2008) make a related demand story but base it on the long-shot bias in horse racing, which, the authors claim, is also relevant for option pricing. This

---

<sup>30</sup> Building on Han (2008), who studied investor sentiment and option prices, Andreou, Kagkadis, and Philip (2014) show that measures of investor sentiment are related to risk-neutral skewness, which in turn relates to the steepness of the index smile.

bias leads to investors overpaying for long-shot bets (such as out-of-the-money calls and puts) compared to safe bets (such as in-the-money calls and puts). They paper argues that this bias could lead to the skew pattern of the index smile, and the argument then continues as above.

### 5.3.2. Models with probability weighting functions

Kliger and Levy (2009) revert the direction of investigation by starting with the pricing kernel puzzle, using power utility, and backing out the implied physical distribution from the risk-neutral distribution. As a result, the implied physical distribution inherits the left-skewed and leptokurtic shape of the risk-neutral distribution, which is incompatible with the physical distribution derived from bootstrapped past S&P 500 returns. Thus, they introduce a probability weighting function in order to reconcile the implied physical distribution with the bootstrapped distribution.<sup>31</sup> The estimated probability weighting functions<sup>32</sup> are inverse-S-shaped in their sample from 1986-1995. Polkovnichenko and Zhao (2013) repeat that study on more recent data, using power utility with a risk aversion coefficient of two, and, for the physical distribution, using an EGARCH model based on past returns. Their probability weighting functions can be S-shaped (2004-2006) or inverse-S-shaped (during the remaining years from 1996 to 2008). The former suggests that investors overweight probabilities in the center of the distribution and underweight the tails, while the pattern reverses for the latter. It is somewhat puzzling that the pricing kernel puzzle tends to be rather stable through time but yields in this setting very different probability weighting functions. The model also does not

---

<sup>31</sup> Gemmill and Shackleton (2005) sketch out a similar idea in an incomplete working paper.

<sup>32</sup> See Quiggin (1982) and Yaari (1987) for rank dependent utility models and Tversky and Kahneman (1992) for cumulative prospect theory models using such functions. Ingersoll (2017) details aggregation results for economies where investors follow cumulative prospect theory. Baele, Driessen, Londono, and Spalt (2014) concentrate on the variance risk premium under cumulative prospect theory.

account for learning; investors do not pay attention to the fact that the physical distribution, as it is being revealed in realized returns, looks different from the reweighted distribution.

Dierkes (2013) makes a nice point about the lack of identification in Polkovnichenko and Zhao (2013), as the utility function cannot be derived separately from the weighting function. He suggests an intriguing solution by fitting several maturities at the same time. That allows the utility function to be the same for all maturities but the weighting function scales with maturity. Empirically, Dierkes (2013) then finds the weighting function to be inverse-S-shaped and the utility function to be convex-concave around the zero percent return.

Chabi-Yo and Song (2013) confirm the findings of Polkovnichenko and Zhao (2009) and document that the probability weighting functions are heavily time-varying, even if they use the VIX as a conditioning variable. They thus extend the model and apply probability weighting to both the return and volatility dimensions of the index in a two period setting. Using S&P 500 and VIX options, they find inverse-S-shaped probability weighting functions, which are now much more stable in comparison with the single state variable model.

#### **5.4. Ambiguity aversion models**

Here, we propose a novel approach based on the smooth ambiguity aversion model of Klibanoff, Marinacci, and Mukerji (2005). The model nests on the one hand the traditional expected utility setting as the ambiguity aversion approaches ambiguity neutrality and on the other hand the maximin utility approach as the ambiguity aversion goes to infinity.<sup>33</sup> Gollier (2011) already mentions that the pricing kernel puzzle can emerge in a smooth ambiguity aversion setting, although without explicitly deriving the formulas and without detailed examples, which we are providing here.

---

<sup>33</sup> For a survey of ambiguity aversion and its relevance for asset pricing, see Epstein and Schneider (2010). For an alternative formulation of ambiguity aversion through Choquet expected utility, see Bassett, Koenker, and Kordas (2004).

Kang, Kim, Lee (2014) use a different set-up where a representative investor faces a stock price process but is also worried that some worst case stock price process with lower drift might be true. Maximizing the minimal utility under those scenarios leads to a pricing kernel which can exhibit the puzzling increasing behavior, and the paper calibrates the model to S&P 500 options and returns data.

Liu, Pan and Wang (2005) investigate rare-event premia driven by uncertainty aversion in the context of a calibrated equilibrium model and demonstrate that the model generates the volatility smile patterns observed in option prices. It is not clear if their model exhibits the pricing kernel puzzle. Drechsler (2013) extends the former model, and his investor selects among several alternative models, which the investor cannot reject. He finally picks the model that is least favorable in terms of expected utility. The resulting equilibrium generates the skew in implied volatilities and the variance risk premium. Going one step further, it would be interesting to explicitly calibrate the model to option data and see if such economy implies a non-monotonic pricing kernel.

#### 5.4.1. The theoretical pricing kernel under ambiguity aversion

We re-derive our simple economy from Section 1, Equations (3-6) in the setting of Klibanoff, Marinacci, and Mukerji (2005). They assume that there are M of the above economies, each with a probability  $z_j$  of occurring for  $j = 1, \dots, M$ . Our representative investor is thus solving the following problem

$$\begin{aligned} \max_{R_i} \sum_{j=1}^M z_j \phi \left( \sum_{i=1}^N \pi_{ij} U(R_i) \right) \\ s.t. \left( \sum_{i=1}^N q_i R_i \right) / R_f = 1 \end{aligned} \tag{11}$$

where  $\pi_{ij}$  is physical probability of state  $i$  occurring in ambiguity setting  $j$  and  $\phi$  is a utility function across ambiguity settings which operates on the expected utility achieved in each

ambiguity setting. Note that the physical probability of being in state  $i$  ( $p_i$ ) is the sum of  $\pi_{ij} z_j$  across ambiguity settings  $j$ . Introducing the Lagrange multiplier  $\lambda$ , we write the  $N$  first order conditions

$$\sum_{j=1}^M z_j \phi' \left( \sum_{k=1}^N \pi_{kj} U(R_k) \right) \pi_{ij} U'(R_i) - \lambda q_i / R_f = 0 \quad \text{for } i=1, \dots, N \quad (12)$$

and solve for  $\lambda$  by summing the  $N$  equations (12)

$$R_f \left( \sum_{i=1}^N \left[ \sum_{j=1}^M z_j \phi' \left( \sum_{k=1}^N \pi_{kj} U(R_k) \right) \pi_{ij} U'(R_i) \right] \right) = \lambda \quad (13)$$

Substituting  $\lambda$  back into equation (12), we obtain after rearranging

$$\frac{q_i / R_f}{\sum_{j=1}^M \pi_{ij} z_j \phi' \left( \sum_{k=1}^N \pi_{kj} U(R_k) \right)} = \frac{U'(R_i)}{R_f \left( \sum_{i=1}^N \left[ \sum_{j=1}^M \pi_{ij} z_j \phi' \left( \sum_{k=1}^N \pi_{kj} U(R_k) \right) U'(R_i) \right] \right)} \quad (14)$$

We can now solve for the pricing kernel  $m$  state by state by multiplying by the left hand-side

denominator and dividing both sides by  $p_i = \sum_{j=1}^M \pi_{ij} z_j$  which is the total probability that state  $i$

occurs.

$$m_i = \frac{q_i}{R_f \sum_{j=1}^M \pi_{ij} z_j} = \frac{\sum_{j=1}^M \pi_{ij} z_j \phi' \left( \sum_{k=1}^N \pi_{kj} U(R_k) \right)}{R_f \sum_{j=1}^M \pi_{ij} z_j} \frac{U'(R_i)}{\sum_{i=1}^N \left[ \sum_{j=1}^M \pi_{ij} z_j \phi' \left( \sum_{k=1}^N \pi_{kj} U(R_k) \right) U'(R_i) \right]} \quad (15)$$

We can readily interpret the pricing kernel formula in comparison to the simple case without ambiguity.<sup>34</sup> There, the pricing kernel is the ratio of marginal utility and expected marginal utility. In the setting with ambiguity aversion, the pricing kernel is the scaled marginal utility in each state divided by a modified expected marginal utility. We explain the modification of expected marginal utility first and then the scaling of the pricing kernel. For the modified expectation, the probabilities of the expectation ( $\pi_{ij} z_j$ ) are being distorted by the marginal ambiguity utility  $\phi' \left( \sum_{k=1}^N \pi_{kj} U(R_k) \right)$ . The resulting quantities are no longer probabilities, i.e. they will not add to one. Thus, the pricing kernel needs to be scaled in order to correct for the modification. The scaling factor is the fraction in front of the marginal utility term in Equation (15). It turns out to be the ratio of the sum of the probabilities ( $\pi_{ij} z_j$ ), which are again being distorted by  $\phi' \left( \sum_{k=1}^N \pi_{kj} U(R_k) \right)$  and the sum of the probabilities themselves ( $\pi_{ij} z_j$ ).

#### 5.4.2. The pricing kernel puzzle in a model of ambiguity aversion

Here we use Equation (15) with power utility formulations with parameters 2 and 7 for ambiguity aversion and risk aversion, respectively. The following choice for  $U(x)$  satisfies the assumption of Klibanoff, Marinacci, and Mukerji (2005) that two utility values need to be independent of  $\gamma$ , here,  $U(1)=0$  and  $U(2)=1$ .<sup>35</sup>

---

<sup>34</sup> Unfortunately, we cannot easily analyze the derivative of the pricing kernel with respect to returns. The resulting expressions are intractable and cannot be nicely segregated into, say, an income and a substitution effect.

<sup>35</sup> Note that alternatively, one could also use  $U(x) = \frac{x^{1-\gamma} - 1}{1-\gamma}$  with  $\gamma \in (0,1)$  but the above formulation allows for a great range of risk aversion coefficients.

$$\phi(x) = \frac{x^{1-\eta} - 1}{1-\eta}; U(x) = \frac{x^{1-\gamma} - 1}{2^{1-\gamma} - 1} \quad (16)$$

Further, we model the 45-day return being lognormally distributed with an annualized mean of 0.06. There are 300 log return levels from -0.99 to +2.00 in steps of 0.01. The investors are ambiguous with respect to annualized volatility, which is assumed to be lognormally distributed with mean log 0.19 and standard deviation 0.10. There are 81 ambiguity settings ranging from -4 to +4 standard deviations in steps of 0.1 standard deviations.<sup>36</sup>

[Figure 6 about here]

We depict the resulting pricing kernel in Figure 6 and it matches quite nicely the pricing kernel puzzle from a particular day (August 15, 2003) in Jackwerth (2004, p.57, figure 11, Panel A). The physical probability distribution (sum of the probabilities  $\pi_{ij} z_j$ ) has a mean of 0.06, standard deviation of 0.19, skewness of 0.00, and kurtosis of 3.12.

Extending the calibration, we set the volatility to 0.20 fixed and introduce ambiguity aversion across the means. The expected annualized mean is 0.08 with an annual standard deviation of 0.15, 0.20, and 0.25. None of these calibrations generate the desired risk aversion puzzle.

The next extension is to introduce large negative jumps (-0.20 mean and 0.10 standard deviation on a grid of 0.01 from -4 to +4 standard deviations) where the investor exhibits ambiguity aversion across the probability of such jumps occurring. Here, we assume a uniform distribution from 0 to 0.8 in steps of 0.02. The return distribution without crashes is modeled being lognormally distributed with a mean of log 0.08 and a volatility of 0.2. There are 300 log return levels from -0.99 to +2.00 in steps of 0.01. The pricing kernel in Figure 7

---

<sup>36</sup> It is interesting to note that, contrary to common models using power utilities, here it does matter how we specify the utility function U; by not subtracting 1 in the numerator, we do not obtain the pricing kernel puzzle of Figure 6 with these parameters.

exhibits the puzzling behavior yet again, although at a somewhat more subdued amplitude than in Figure 6.

[Figure 7 about here]

Thus, a simple one-period ambiguity aversion model can exhibit the pricing kernel puzzle. Whereas ambiguity over the mean return of the underlying does not introduce the pricing kernel puzzle, ambiguity over either volatility or the probability of large crashes exhibits the empirical pricing kernel puzzle.

## 6. Bounds on option prices

The literature on bounds on option prices takes a different perspective on the pricing kernel puzzle. The pricing kernel puzzle is about analyzing the empirical pricing kernel, given risk-neutral and physical distributions, where the pricing kernel turns out to be non-decreasing in returns. Turning the problem around, one can ask what are the highest and lowest option prices still compatible with a monotonically decreasing pricing kernel? This approach was developed in Perrakis and Ryan (1984) with the restrictions that the pricing kernel has to be positive, decreasing, and that it prices the stock and the bond and one reference option traded in the market. The resulting linear program then looks (for a call option with given strike price  $K$ ) as follows:

$$\begin{aligned}
 & \text{Max/Min } E[m (R^*S - K)^+] \\
 \text{s.t. } & E[m] = B \\
 & E[m R^*S] = S \\
 & E[m (R^*S - K^0)^+] = C(K^0) \\
 & m > 0, m \text{ decreasing in wealth } R^*S
 \end{aligned} \tag{17}$$

where  $S$  is the initial stock price,  $B$  the unit bond price, and  $K^0$  the strike price of the observed reference option. Dividends are assumed to be zero for ease of exposition. A long literature

ensued which extends the above linear program approach, adding e.g. bid/ask spreads and transaction costs, see the survey of Constantinides, Jackwerth, and Perrakis (2008).

The resulting bounds are driven by pricing kernels which tend to be extreme; exhibiting steep drops after almost flat sections. Cochrane and Saa-Requejo (2000) address the problem of such unrealistic pricing kernels. They essentially work within the above set-up while also restricting the volatility of the pricing kernel, which leads to smoother pricing kernels and tighter bounds. Bernardo and Ledoit (2000) offer an alternative restriction by limiting the ratio of expected gains and expected losses of a security; ruling out that securities are priced much too low or high compared to their fair value.<sup>37</sup> Pyo (2011) achieves this goal by the ad-hoc restriction that price deviations are limited by deviations of observed prices from model prices based on a predetermined (power) pricing kernel. His approach seems to boil down to restricting the pricing kernel to stem from a power utility while restricting the risk aversion coefficient.

The earlier papers solved the linear program explicitly and were thus limited in the complexity of the linear program, e.g., they could only handle one reference option and extending it to two was already a difficult task. Relying simply on computer solutions to the linear program, Constantinides, Jackwerth, and Perrakis (2009) can compute bounds for S&P 500 index options while taking into account all observed options as reference assets and even formulating the linear program over two steps instead of one. Further, they use the analytical bounds from Constantinides and Perrakis (2002) for continuous, intermediate trading and proportional transaction costs. Empirically, they find a substantial number of options to be located outside their bounds, consistent with the pricing kernel puzzle. Wallmeier (2015)

---

<sup>37</sup> Marroquin-Martinez and Moreno (2013) extends Cochrane and Saa-Requejo (2000) and Bernardo and Ledoit (2000) to settings with stochastic volatility and finds the resulting bounds to be tighter than in the original papers.

replicates their work and finds far fewer violations. This difference comes about as Wallmeier (2015) uses option implied information from just hours ago to adjust the physical distribution, while Constantinides, Jackwerth, and Perrakis (2009) rely on information further into the past. The concern is that using more recent option implied information will eventually move the physical distribution so close to the risk-neutral, that one can no longer detect bound violations.

More interesting is the question if, using the earlier information of Constantinides, Jackwerth, and Perrakis (2009), one can profitably trade based on bound violations. That exercise can be found in Constantinides, Czerwonko, Jackwerth, and Perrakis (2011), now using options on futures on S&P 500 and employing the analytical bounds of Constantinides and Perrakis (2007), which are suitable for these American options. The results suggest that trading strategies involving out-of-bounds options are superior to pure stock-and-bond strategies for all risk-averse investors.

## **7. Conclusion and outlook**

In our survey of the pricing kernel puzzle, we recount the history, starting with the canonical papers which around the year 2000 divided risk-neutral distributions of S&P 500 returns by the physical distributions. These empirical pricing kernels exhibited increasing sections, which are inconsistent with simple representative investor models with a single state variable. Evidence from indexes in other countries and other periods finds the same puzzling behavior. We also discuss the (sparse) literature, which cannot detect the pricing kernel puzzle in the data and try to understand the reasons.

A number of statistical tests suggest the presence of the pricing kernel puzzle. What is still missing is a critical analysis and comparison of the several tests, which so far exist. Are their

assumptions realistic? Are certain tests better than others? There is still no agreement on which test to use as the standard test of the pricing kernel puzzle.

Much room is given to the potential explanations of the pricing kernel puzzle, starting with simple one-state-variable formulations and then moving to more complex settings. Similarly to the tests, many of the solutions are stand-alone model with little empirical validation. Mostly, they concentrate on a calibration, which, using some stylized facts, exhibits the pricing kernel puzzle. Much work is still needed in sorting through the alternative models and grading them according to their compatibility with the data. Ideally, some of the solutions might be joined in a nested model, allowing for a proper test of the different features. It would be interesting to know more about the true mechanism of what drives the pricing kernel puzzle. Interesting research along those lines is trying to explain the time-series patterns of the pricing kernel puzzle (e.g., its severity) using explanatory variables. This challenging work is still in its infancy and, as of now, still underwhelming.

For a glimpse into the future of the pricing kernel puzzle, one might want to consider the bivariate estimation of risk-neutral and physical distributions in Jackwerth and Vilkov (2015). Those bivariate risk-neutral distributions can normally only be obtained with the help of options written on both assets simultaneously, but Jackwerth and Vilkov (2015) were able to achieve this feat in the dimensions of index returns and volatility employing longer-dated options on the S&P 500. Dividing the two distributions into each other allows one to extract for the first time a bivariate pricing kernel.

## Bibliography

- Ait-Sahalia, Yacine, and Andrew W. Lo, 1998, Nonparametric Estimation of State-Price Densities Implicit in Financial Asset Prices, *Journal of Finance*, 53, 499-547.
- Ait-Sahalia, Yacine, and Andrew W. Lo, 2000, Nonparametric Risk Management and Implied Risk Aversion, *Journal of Econometrics*, 94, 9–51.
- Ait-Sahalia, Yacine, and Jefferson Duarte, 2003, “Nonparametric Option Pricing Under Shape Restrictions,” *Journal of Econometrics*, 116, 9-47.
- Ait-Sahalia, Yacine, and Michael Brandt, 2007, “Consumption and Portfolio Choice with Option-Implied State Prices,” Working Paper, Princeton University.
- Alcock, Jamie, and Godfrey Smith, 2014, “Testing Alternative Measure Changes in Nonparametric Pricing and Hedging of European Options,” *Journal of Futures Markets*, 34, 320-345.
- Andersen, Steffen, John Fountain, Glenn W. Harrison, and E. Elisabet Rutstroem, 2014, “Estimating Subjective Probabilities,” *Journal of Risk and Uncertainty*, Vol 48, 207-229.
- Andreoni, James, and Charles Sprenger, 2012a, “Estimating Time Preferences from Convex Budgets,” *American Economic Review*, 102, 3333-3356.
- Andreoni, James, and Charles Sprenger, 2012b, “Risk Preferences are not Time Preferences,” *American Economic Review*, 102, 3357-3376.
- Andreou, Panayiotis C., Anastasios Kagkadis, and Dennis Philip, 2014, “Investor Sentiments, Rational Beliefs, and Option Prices,” Working paper, Lancaster University.
- Audrino, Francesco, and Pirmin Meier, 2012, “Empirical Pricing Kernel Estimation Using a Functional Gradient Descent Algorithm Based on Splines,” Working paper, University of St. Gallen.
- Audrino, Francesco, R. Huitema, and M. Ludwig. 2015, “An Empirical Analysis of the Ross Recovery,” Working paper, University of St. Gallen.
- Babaoglu, Kadir, Peter Christoffersen, Steven Heston, and Kris Jacobs, 2016, “Option Valuation with Volatility Components, Fat Tails, and Non-Monotonic Pricing Kernels,” Working paper, University of Toronto.
- Backus, David, Mikhail Chernov, and Ian Martin, 2011, “Disasters Implied by Equity Index Options,” *Journal of Finance*, 66, 1969-2012.
- Baele, Lieven, Joost Driessen, Juan M. Londono, and Oliver G. Spalt, 2014, “Cumulative Prospect Theory and the Variance Premium,” Working paper, Federal Reserve Board of Governors.
- Bahra, Bhupinder, 1997, “Implied Risk-Neutral Probability Density Functions from Option Prices: Theory and Application,” Working paper, Bank of England, no. 66.
- Bakshi, G., N. Kapadia, and D. Madan, 2003, “Stock Return Characteristics, Skew Laws, and Differential Pricing of Individual Equity Options,” *Review of Financial Studies*, 16, 101–143.
- Bakshi, Gurdip, and Dilip B. Madan, 2007, “Investor Heterogeneity, Aggregation, and the Non-Monotonicity of the Aggregate Marginal Rate of Substitution in the Price of Market-Equity,” Working paper, University of Maryland.
- Bakshi, Gurdip, and George Panayotov, 2008, “A Framework for Studying Option Mispricing: A General Test and Empirical Evidence,” Working paper, University of Maryland.
- Bakshi, Gurdip, Dilip B. Madan and George Panayotov, 2010, “Returns of Claims on the Upside and the Viability of U-Shaped Pricing Kernels,” *Journal of Financial Economics*, 97, 130-154.

- Bakshi, Gurdip, Dilip B. Madan and George Panayotov, 2014, "Heterogeneity in Beliefs and Volatility Tail Behavior," *Journal of Financial and Quantitative Analysis*, forthcoming.
- Bali, Turan G., Nusret Cakici, Fousseni Chabi-Yo, and Scott Murray, 2014, "Risk-Neutral Moments of Option Returns," Working paper, Georgetown University.
- Bansal, Ravi, and Amir Yaron, 2004, "Risks for the Long Run: A Potential Resolution of Asset Pricing Puzzles," *Journal of Finance*, 59, 1481-1509.
- Barone-Adesi, Giovanni, and Hakim Dall'O, 2010, "Is the Price Kernel Monotone?," Working paper, Swiss Finance Institute.
- Barone-Adesi, Giovanni, Lorian Mancini, and Hersh Shefrin, 2012, "Behavioral Finance and the Pricing Kernel Puzzle: Estimating Sentiment, Risk Aversion, and Time Preference," Working paper, Swiss Finance Institute.
- Barone-Adesi, Giovanni, Robert F. Engle and Lorian Mancini, 2008, "A GARCH Option Pricing Model with Filtered Historical Simulation," *Review of Financial Studies*, 21, 1223-1258.
- Barsky, Robert B., F. Thomas Juster, Miles S. Kimball, and Matthew D. Shapiro, 1997, "Preference Parameters and Behavioral Heterogeneity: An Experimental Approach in the Health and Retirement Study," *Quarterly Journal of Economics* 112, 537-579.
- Bartunek, K. S. and M. Chowdhury, 1997, "Implied Risk Aversion Parameter from Option Prices," *Financial Review*, 32, 107-124.
- Bassett, Gilbert W., Roger Koenker, and Gregory Kordas, 2004, "Pessimistic Portfolio Allocation and Choquet Expected Utility," *Journal of Financial Econometrics*, 2, 477-492.
- Bates, D., 2000, "Post-'87 Crash Fears in S&P 500 Futures Options," *Journal of Econometrics*, 94, 181-238.
- Bates, D., 2012, "U.S. Stock Market Crash Risk, 1926-2010," *Journal of Financial Economics*, 105, 229-259.
- Bates, David S., 1996a, "Testing Option Pricing Models," Chapter 20 of G.S. Maddala and C. R. Rao, editors, *Statistical Methods in Finance (Handbook of Statistics, Vol. 14)*, pp.567-611.
- Bates, David S., 1996b, "Jumps and Stochastic Volatility: Exchange Rate Processes Implicit in Deutsche Mark Options," *Review of Financial Studies*, 9, 69-107.
- Bates, David S., 2008, "The Market for Crash Risk," *Journal of Economic Dynamics and Control*, 32, 2291-2321.
- Beare, Brendan K., 2011, "Measure Preserving Derivatives and the Pricing Kernel Puzzle," *Journal of Mathematical Economics*, 47, 689-697.
- Beare, Brendan K., and L. Schmidt, 2015, "An Empirical Test of Pricing Kernel Monotonicity," *Journal of Applied Econometrics*, forthcoming.
- Beare, Brendan K., and L. Schmidt, 2015, "Empirical implications of the pricing kernel puzzle for the return on contingent claims," *Management and Science*, Working Paper.
- Belomestny, Denis, Shujie Ma, and Wolfgang Haerdle, 2015, "Sieve estimation of the minimal entropy martingale marginal density with application to pricing kernel estimation," Working Paper, Humboldt University.
- Benth, Fred Espen, Martin Groth, and Carl Lindberg, 2010, "The implied risk aversion from utility indifference option pricing in a stochastic volatility model," *International Journal of Applied Mathematics & Statistics*, 16, 11-37.
- Benzoni, Luca, Pierre Collin-Dufresne and Robert S. Goldstein, 2011, "Explaining Asset Pricing Puzzles Associated with the 1987 Market Crash," *Journal of Financial Economics*, 101, 552-573.

- Bernardo, Antonio E., and Olivier Ledoit, 2000, "Gain, Loss, and Asset Pricing," *Journal of Political Economy*, 108, 144–172.
- Bliss, Robert R., and Nikolaos Panigirtzoglou, 2002, "Testing the Stability of Implied Probability Density Functions," *Journal of Banking and Finance*, 26, 381-422.
- Bliss, Robert R., and Nikolaos Panigirtzoglou, 2004, "Option-Implied Risk Aversion Estimates," *Journal of Finance*, 59, 407-446.
- Boes, Mark-Jan, Feike C. Drost, and Bas J. M. Werker, 2007, "Nonparametric Risk-Neutral Joint Return and Volatility Distributions," Working paper, VU University Amsterdam and Tilburg University.
- Bollen, Nicholas P. B., and Robert E. Whaley, 2004, "Does Net Buying Pressure Affect the shape of Implied Volatility Functions?," *Journal of Finance*, 59, 711-754.
- Bollerslev, Tim, and Viktor Todorov, 2011, "Tails, Fears and Risk Premia," *Journal of Finance*, 66, 2165-2211.
- Borovicka, Jaroslav, Lars Peter Hansen, and Jose A. Scheinkman, 2015, "Misspecified Recovery," Working Paper, NBER.
- Branger, Nicole, Alexandra Hansis, and Christian Schlag, 2011, "The Impact of Risk Premia on Expected Option Returns," Working paper, University of Muenster.
- Branger, Nicole, and Christian Schlag, 2004, "Why is the Index Smile So Steep?," *Review of Finance*, 8, 109-127.
- Breeden, Douglas T., and Robert H. Litzenberger, 1978, Prices of State-Contingent Claims Implicit in Options Prices, *Journal of Business* 51, 621–651.
- Broadie, Mark, Mikhail Chernov, and Michael Johannes, 2009, "Understanding Index Option Returns," *Review of Financial Studies*, 22, 4493-4529.
- Brown, David P. and Jens Carsten Jackwerth, 2012, "The Pricing Kernel Puzzle: Reconciling Index Option Data and Economic Theory," in: Contemporary Studies in Economics and Financial Analysis: Derivative Securities Pricing and Modelling, Vol. 94, 155-183, eds. J. A. Batten and N. Wagner, Emerald Group, Bingley, UK, 2012.
- Buraschi, Andrea and Jens C. Jackwerth, 2001, "The Price of a Smile: Hedging and Spanning in Option Markets," *Review of Financial Studies*, 14, 495-527.
- Campbell, J., and J. Cochrane, 1999, "By Force of Habit: A Consumption-Based Explanation of Aggregate Stock Market Behavior," *Journal of Political Economy*, 107, 205–251.
- Carr, Peter and Jiming Yu, 2012, "Risk, Return, and Ross Recovery," *Journal of Derivatives*, 20, 38-59.
- Carr, Peter, Hélyette Geman, Dilip P. Madan, and Marc Yor, 2002, "The Fine Structure of Asset Returns: An Empirical Investigation," *Journal of Business*, 75, 305-332.
- Chabi-Yo, Fousseni, 2012, "Pricing Kernels with Stochastic Skewness and Volatility Risk," *Management Science*, 58, 624-640.
- Chabi-Yo, Fousseni, and Zhaogang Song, 2013, "Recovering the Probability Weights of Tail Events with Volatility Risk from Option Prices," Working Paper, Ohio State.
- Chabi-Yo, Fousseni, René Garcia, and Eric Renault, 2008, "State Dependence can Explain the Risk Aversion Puzzle," *Review of Financial Studies*, 21, 973-1011.
- Chaudhuri, Ranadeb, and Mark D. Schroder, 2015, "Monotonicity of the Stochastic Discount Factor and Expected Option Returns," *Review of Financial Studies*, 28, 1462-1505.
- Chernov, Mikhail, 2003, "Empirical Reverse Engineering of the Pricing Kernel," *Journal of Econometrics*, 116, 329-364.
- Chernov, Mikhail, and Eric Ghysels, 2000, "A Study Towards a Unified Approach to the Joint Estimation of Objective and Risk-Neutral Measures for the Purpose of Options Valuation," *Journal of Financial Economics*, 56, 207-458.
- Chetty, Raj, and Adam Szeidl, 2007, "Consumption Commitments and Risk Preferences," *Quarterly Journal of Economics*, 122, 831-877.

- Christoffersen, Peter, Steven Heston, and Kris Jacobs, 2013, "Capturing Option Anomalies with a Variance-Dependent Pricing Kernel," *Review of Financial Studies*, 26, 1962-2006.
- Christoffersen, Peter, Steven Heston, and Kris Jacobs, 2009, "The Shape and Term Structure of the Index Option Smirk: Why Multifactor Stochastic Volatility Models Work So Well," *Management Science*, 55, 1914-1932.
- Cochrane, John H., and Jesus Saa-Requejo, 2000, "Beyond Arbitrage: Good-Deal Asset Price Bounds in Incomplete Markets," *Journal of Political Economy*, 108, 79-119.
- Constantinides, George M., and Stylianos Perrakis, 2002, "Stochastic Dominance Bounds on Derivative Prices in a Multiperiod Economy with Proportional Transaction Costs," *Journal of Economic Dynamics and Control*, 26, 1323-1352.
- Constantinides, George M., and Stylianos Perrakis. 2007, "Stochastic Dominance Bounds on American Option Prices in Markets with Frictions," *Review of Finance*, 11, 71-115.
- Constantinides, George M., Jens Carsten Jackwerth and Alexi Savov, 2013, "The Puzzle of Index Option Returns," *Review of Asset Pricing Studies*, 3, 229-257.
- Constantinides, George M., Jens Carsten Jackwerth and Stylianos Perrakis, 2008, "Option Pricing: Real and Risk-Neutral Distributions," in: *Handbooks in Operations Research and Management Science: Financial Engineering*, Vol. 15, 565-591, eds. J.R. Birge and V. Linetsky, Elsevier, Amsterdam, 2007.
- Constantinides, George M., Jens Carsten Jackwerth, and Stylianos Perrakis, 2009, "Mispricing of S&P 500 Index Options," *Review of Financial Studies*, 22, 1247-1277.
- Constantinides, George M., Michal Czerwonko, Jens Carsten Jackwerth and Stylianos Perrakis, 2011, "Are Options on Index Futures Profitable for Risk Averse Investors? Empirical Evidence," *Journal of Finance*, 66, 1401-1431.
- Coutant, Sophie, 1999, "Implied Risk Aversion in Option Prices Using Hermite Polynomials," Working paper, Banque de France.
- Coutant, Sophie, 2000, "Time-Varying Implied Risk Aversion in Option Prices Using Hermite Polynomials," Working paper, Banque de France.
- Coval, Joshua D., and Tyler Shumway, 2001, "Expected Option Returns," *Journal of Finance*, 56, 983-1009.
- Cuesdeanu, Horatio, and Jens Jackwerth, 2016, "The Pricing Kernel Puzzle in Forward Looking Data," Working paper, University of Konstanz.
- David, Alexander, and Pietro Veronesi, 2011, "Investor and Central Bank Uncertainty and Fear Measures Embedded in Index Options," NBER Working Paper No. 16764.
- Detlefsen, Kai, Wolfgang Haerdle, and Rouslan Moro, 2010, "Empirical Pricing Kernels and Investor Preferences," *Journal of Mathematical Methods in Economics and Finance*, 3, 19-48.
- Diebold, Francis X., Anthony S. Tay, and Kenneth F. Wallis, 1999, Evaluating Density Forecasts of Inflation: The Survey of Professional Forecasters, in Robert F. Engle, and Halbert White, eds. *Cointegration, Causality, and Forecasting: A Festschrift in Honour of Clive W. J. Granger* (Oxford University Press, New York), 76-90.
- Diebold, Francis X., Todd A. Gunther, and Anthony S Tay, 1998, Evaluating Density Forecasts with Applications to Financial Risk Management, *International Economic Review*, 39, 863-883.
- Dierkes, Maik, 2013, "Probability Weighting and Asset Prices," Working paper, University of Muenster.
- Dittmar, Robert F., 2002, "Nonlinear Pricing Kernels, Kurtosis Preference, and Evidence from the Cross Section of Equity Returns," *Journal of Finance*, 57, 369-403.
- Drechsler, Itamar, 2013, "Uncertainty, Time-Varying Fear, and Asset Prices," *Journal of Finance*, 68, 1843-1889.

- Duan, Jin-Chuan, and Weiqi Zhang, 2014, "Forward-Looking Market Risk Premium," *Management Science*, 60, 521-538.
- Dybvig, Philip H, 1988, "Distributional Analysis of Portfolio Choice," *Journal of Business*, 61, 369-93.
- Epstein, Larry G., and Martin Schneider, 2010, "Ambiguity and Asset Markets," *Annual Review of Financial Economics*, 2, 315-346.
- Epstein, Larry G., and Stanley Zin, 1989, „Substitution, Risk Aversion, and the Temporal Behavior of Consumption and Asset Returns: A Theoretical Framework,“ *Econometrica* 57, 937–969.
- EZB, 2011, "Der Informationsgehalt von Optionspreisen während der Finanzkrise," Monatsbericht, EZB Februar.
- Fengler, Matthias, and Lin-Yee Hin, 2015, "Semi-Nonparametric Estimation of the Call Price Surface Under No-Arbitrage Constraints," *Journal of Econometrics*, 184, 242-261.
- Figlewski, Stephen, 2010, "Estimating the Implied Risk-Neutral Density for the U.S. Market Portfolio," in T. Bollerslev, J. Russell and M. Watson (Eds), *Volatility and Time Series Econometrics: Essays in Honor of Robert F. Engle*, Oxford University Press, Oxford and New York, pp. 323–353.
- Figlewski, Stephen, and Muhammad Fahd Malik, 2014, "Options on Leveraged ETFs: A Window on Investor Heterogeneity," Working Paper, New York University.
- Friedman, Milton, and L. J. Savage, 1948. "Utility Analysis of Choices Involving Risk," *Journal of Political Economy*, 56, 279–304.
- Garcia, René, Richard Luger, and Eric Renault, 2003, "Empirical Assessment of an Intertemporal Option Pricing Model with Latent Variables," *Journal of Econometrics*, 116, 49-83.
- Garleanu, Nicolae, Lasse H. Pedersen, and Allen M. Poteshman, 2009, "Demand-Based Option Pricing," *Review of Financial Studies*, 22, 4259-4299.
- Gemmill, Gordon, and Mark B. Shackleton, 2005, "Prospect Theory and Option Prices: Evidence from S&P 500 Index Options," Working paper, Warwick University.
- Giacomini, Enzo, and Wolfgang Haerdle, 2008, "Dynamic Semiparametric Factor Models in Pricing Kernel Estimation," in: *Functional and Operational Statistics*, Dabo-Niang, S. and Ferraty, F. (Eds), Contributions to Statistics, Springer Verlag, 181-187.
- Giacomini, Enzo, Michael Handel, and Wolfgang Haerdle, 2008, "Time Dependent Relative Risk Aversion," in: *Risk Assessment: Decisions in Banking and Finance*, Heidelberg, Physica-Verlag, 15-46.
- Glosten, L. R., R. Jagannathan, and D. E. Runkle, 1993, "On the Relation between the Expected Value and the Volatility of the Nominal Excess Return on Stocks," *Journal of Finance*, 48, 1779–1801.
- Gollier, Christian, 2011, "Portfolio Choices and Asset Prices: The Comparative Statics of Ambiguity Aversion," *Review of Economic Studies*, 78, 1329-1344.
- Golubev, Yuri, Wolfgang Haerdle, and Roman Timofeev, 2014, "Testing Monotonicity of Pricing Kernels," in: *AStA Advances in Statistical Analysis*, Springer, Volume 98, 305-326
- Grith, Maria, Wolfgang Haerdle, and Juhyun Park, 2013 "Shape Invariant Modeling Pricing Kernels and Risk Aversion," *Journal of Financial Econometrics*, 11, 370-399.
- Grith, Maria, Wolfgang Haerdle, and Volker Kraetschmer, 2015, "Reference Dependent Preferences and the EPK Puzzle," *Review of Finance*, forthcoming.
- Haas, Jose Renato, Jose Santiago Fajardo Barbachan, and Aquiles Rocha de Farias, 2011, "Estimating Relative Risk Aversion, Risk-Neutral and Real-World Densities Using Brazilian Real Currency Options," Working Paper, Banco Central do Brasil.
- Haerdle, Wolfgang, Maria Grith and Andrija Mihoci, 2014, "Cross Country Evidence for the EPK Paradox," Working Paper, Humboldt University.

- Haerdle, Wolfgang, and Volker Kraetschmer, and Rouslan Moro, 2009, "A Microeconomic Explanation of the EPK Paradox," Working paper, Humboldt University.
- Haerdle, Wolfgang, Yarema Okhrin, and Weining Wang, 2015, "Uniform Confidence Bands for Pricing Kernels," *Journal of Financial Econometrics*, 13, 376-413.
- Haley, M. R., and T. Walker, 2010, "Alternative tilts for nonparametric option pricing," *Journal of Futures Markets*, 30, 983-1006.
- Han, Bing, 2008, "Investor Sentiment and Option Prices," *Review of Financial Studies*, 21, 387-414.
- Han, Qian, and Calum G. Turvey, 2010, "Equilibrium Market Prices for Risks and Market Risk Aversion in a Complete Stochastic Volatility Model with Habit Formation," Working Paper, Cornell University.
- Haug, Jorgen, Thorsten Hens, and Peter Woehrmann, 2013, "Risk Aversion in the Large and in the Small," *Economics Letters*, 118, 310-313.
- Hens, Thorsten, and Christian Reichlin, 2013, "Three Solutions to the Pricing Kernel Puzzle," *Review of Finance*, 17, 1065-1098.
- Heston, S. L., 1993, "A Closed-Form Solution for Options with Stochastic Volatility with Applications to Bond and Currency Options," *Review of Financial Studies*, 6, 327-343.
- Heston, S., and S. Nandi. 2000, "A Closed-Form GARCH Option Pricing Model," *Review of Financial Studies*, 13, 585-626.
- Hill, Simon I., 2013, "Tracking Value at Risk through Derivative Prices," *Journal of Computational Finance*, 16, 79-122.
- Hodges, Stewart D., Robert G. Tompkins, and William T. Ziemba, 2008, "The Favorite-Longshot Bias in S&P 500 and FTSE 100 Index Futures Options: The Return to Bets and the Cost of Insurance," in Hausch, Donald B. and William T. Ziemba, 2008, *Handbook of Sports and Lottery Markets*, Elsevier, Amsterdam.
- Huang, Darien, and Ivan Shaliastovich, 2014, "Risk Adjustment and the Temporal Resolution of Uncertainty: Evidence from Options Markets." Working Paper, University of Pennsylvania.
- Ingersoll, Jonathan E., 2017, "Cumulative Prospect Theory, Aggregation, and Pricing," *Critical Finance Review*, forthcoming.
- Jackwerth, Jens Carsten, 1999, "Option-Implied Risk-Neutral Distributions and Implied Binomial Trees: A Literature Review," *Journal of Derivatives*, 2, 66-82.
- Jackwerth, Jens Carsten, 2000, "Recovering Risk Aversion from Option Prices and Realized Returns," *Review of Financial Studies* 13, 433-467.
- Jackwerth, Jens Carsten, 2004, "Option-Implied Risk-Neutral Distributions and Risk-Aversion," *The Research Foundation of AIMR*, Charlottesville, USA, 2004.
- Jackwerth, Jens Carsten, and Grigory Vilkov, 2015, "Asymmetric Volatility Risk: Evidence from Option Markets," Working paper, University of Konstanz.
- Jackwerth, Jens Carsten, and Mark Rubinstein, 1996, "Recovering Probability Distributions from Option Prices," *Journal of Finance*, 51, 1611-1631.
- Jackwerth, Jens Carsten, and Marco Menner, 2015, "Does the Ross Recovery Theorem Work Empirically?," Working paper, University of Konstanz.
- Jensen, Christian S., David Lando, and Lasse H. Pedersen, 2016, "Generalized Recovery", Working Paper, Copenhagen Business School.
- Kang, Byung Jin, and Tong Suk Kim, 2006, "Option-implied Risk Preferences: An Extension to Wider Classes of Utility Functions," *Journal of Financial Markets*, 9, 180-198.
- Kang, Byung Jin, Tong Suk Kim, and Hyo Seob Lee, 2010, "Option-Implied Preferences with Model Uncertainty," *Journal of Futures Markets*, 34, 6, 498-515.
- Kelly, Bryan, Lubos Pastor, and Pietro Veronesi, 2015, "The Price of Political Uncertainty: Theory and Evidence from the Option Market," *Journal of Finance*, forthcoming.

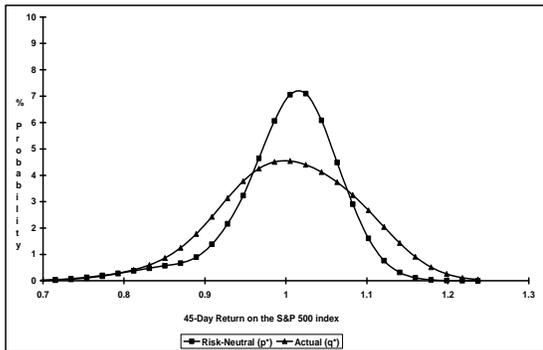
- Kiesel, Ruediger, and Florentin Rahe, 2015, "Option pricing under time-varying risk-aversion with applications to risk forecasting," Working paper, University of Duisburg-Essen.
- Kitsul, Yuriy, and Jonathan H. Wright, 2013, "The Economics of Options-Implied Inflation Probability Density Functions", *Journal of Financial Economics*, 110, 696–711.
- Klibanoff, Peter, Massimo Marinacci and Sujoy Mukerji, 2005, "A Smooth Model of Decision Making Under Ambiguity," *Econometrica*, 73, 1849-1892.
- Kliger, Doron, and Ori Levy, 2009, "Theories of choice under risk: Insights from financial markets," *Journal of Economic Behavior and Organization*, 71, 330-346.
- Kocherlakota, Narayana R., 1996, "The Equity Premium: It's Still a Puzzle," *Journal of Economic Literature*, 34, 42–71.
- Kostakis, Alexandros, Nikolaos Panigirtzoglou, and George Skiadopoulos, 2011, "Market Timing with Option-Implied Distributions: A Forward-Looking Approach," *Management Science*, 57, 1231-1249.
- Laury, Susan K, Melayne Morgan McInnes, J Todd Swarthout, and Erica Von Nessen, 2012, "Avoiding the Curves: Direct Elicitation of Time Preferences", *Journal of Risk and Uncertainty*, 44, 181-217.
- Leisen, Dietmar, 2014, "Small Sample Bias in Pricing Kernel Estimations," Working paper, University of Mainz.
- Li, Haitao, and Feng Zhao, 2009, "Nonparametric Estimation of State-Price Densities Implicit in Interest Rate Cap Prices," *Review of Financial Studies*, 22, 4335-4376.
- Linn, Matthew, Sophie Shive and Tyler Shumway, 2014, "Pricing Kernel Monotonicity and Conditional Information," Working paper, University of Michigan.
- Lioui, Abraham, and Ron Malka, 2004, "Revealing the Parameter of Risk-Aversion From Option Prices When Markets Are Incomplete: Theory and Evidence," Working Paper, Bar Ilan University.
- Liu, Jun, Jun Pan, and Tan Wang, 2005, An Equilibrium Model of Rare-Event Premia and Its Implication for Option Smirks, *The Review of Financial Studies*, 18, 131-164.
- Liu, Xiaoquan, Jing-Ming Kuo, and Jerry Coakley, 2015, "A Pricing Kernel Approach to Valuing Options on Interest Rate Futures," *European Journal of Finance*, 21, 93-110.
- Liu, Xiaoquan, Mark B. Shackleton, Stephen J. Taylor and Xinzhong Xu, 2009, "Empirical Pricing Kernels Obtained from the UK Index Options Market," *Applied Economics Letters*, 16, 989-993.
- Liu, Xiaoquan, Mark B. Shackleton, Stephen J. Taylor and Xinzhong Xu, 2007, "Closed-form Transformations from Risk-neutral to Real-world Distributions," *Journal of Banking and Finance*, 31, 1501-1520.
- Ludwig, Markus, 2015, "Robust Estimation of Shape-Constrained State Price Density Surfaces," *Journal of Derivatives*, 22, 56-72.
- Lundtofte, Frederik, 2010, "Implied Volatility and Risk Aversion in a Simple Model with Uncertain Growth," *Economics Bulletin*, 30, 182-191.
- Madan, Dilip, and F. Milne, 1994, "Contingent Claims Valued and Hedged by Pricing and Investing in a Basis," *Mathematical Finance*, 4, 223–245.
- Marroquin-Martinez, Naroa and Manuel Moreno, 2013, "Optimizing bounds on security prices in incomplete markets. Does stochastic volatility specification matter?" *European Journal of Operational Research*, 225, 429-442.
- Mehra, Rajnish and Edward C. Prescott, 1985, "The Equity Premium: A Puzzle," *Journal of Monetary Economics*, 15, 145-161.
- Mehra, Rajnish, 2006, "The Equity Premium Puzzle: A Review," *Foundations and Trends in Finance*, 2, 1-81.
- Melick, William R. and Charles P. Thomas, 1997, "Recovering an Asset's Implied PDF from Option Prices: An Application to Crude Oil during the Gulf Crisis," *Journal of Financial and Quantitative Analysis*, 32, 91-115.

- Merton, Robert C., 1976, "Option Pricing When Underlying Stock Returns are Discontinuous," *Journal of Financial Economics*, 3, 125–144.
- Meyer, Donald, and Jack Meyer, 2005a, "Relative Risk Aversion: What Do We Know?" *Journal of Risk and Uncertainty*, 31, 243-262.
- Meyer, Donald, and Jack Meyer, 2005b, "Risk Preferences in Multi-period Consumption Models, the Equity Premium Puzzle, and Habit Formation Utility," *Journal of Monetary Economics*, 52, 1497-1515.
- Ni, Sophie Xiaoyan, 2009, "Stock Option Returns: A Puzzle," Working paper, Hong Kong University of Science and Technology.
- Overberg, Sonja Talea, 2009, "Empirical Estimation of Risk Aversion from Surveys," Working paper, University of Konstanz.
- Pan, Jun, 2002, "The Jump-Risk Premia Implicit in Options: Evidence from an Integrated Time-Series Study," *Journal of Financial Economics*, 63, 3-50.
- Patton, Andrew J. and Allan Timmermann, 2010, "Monotonicity in Asset Returns: New Tests with Applications to the Term Structure, the CAPM, and Portfolio Sorts," *Journal of Financial Economics*, 98, 605-625.
- Perignon, Christophe and Christophe Villa, 2002, "Extracting Information from Options Markets: Smiles, State-Price Densities and Risk Aversion," *European Financial Management*, 8, 495-513.
- Perrakis, Stelios, and P. Ryan, 1984, "Option Pricing Bounds in Discrete Time", *Journal of Finance*, 39, 2, 519-525.
- Polkovnichenko, Valery, and Feng Zhao, 2013, "Probability Weighting Functions Implied by Options Prices," *Journal of Financial Economics*, 107, 580-609.
- Pozzi, L. C. G., C. G. de Vries and J. Zenhorst, 2010, "World Equity Premium Based Risk Aversion Estimates," Tinbergen Institute Discussion Paper 2010-007/2, Tinbergen Institute.
- Pyo, Unyong, 2011, "Minmax Price Bounds in Incomplete Markets," *Journal of Economics and Finance*, 35, 3, 274-295.
- Quiggin, J., 1982, "A Theory of Anticipated Utility," *Journal of Economic and Behavioral Organization*, 3, 323–343.
- Rieger, Marc O., 2011, "Co-monotonicity of optimal investments and the design of structured financial products," *Mathematics and Statistics*, 15, 27-55.
- Rosenberg, Joshua V., and Robert F. Engle, 2002, "Empirical pricing kernels," *Journal of Financial Economics* 64, 341–372.
- Ross, Stephen, 2015, "The Recovery Theorem," *Journal of Finance*, 70, 615–648.
- Rubinstein, Mark, 1994, "Implied Binomial Trees," *Journal of Finance*, 49, 771-818.
- Sala, Carlo, and Giovanni Barone-Adesi, 2015, "The Pricing Kernel Anomaly: the Case of the Information that did not Bark" Working paper, Swiss Finance Institute.
- Santa-Clara, Pedro, and Shu Yan, 2010, "Crashes, Volatility, and the equity premium: Lessons from S&P 500 Options," *Review of Economics and Statistics*, 92, 435-451.
- Sarantopoulou-Chioureaa, Sylvia and George Skiadopoulos, 2015, "A New Predictor of Real Economic Activity: The S&P 500 Option Implied Risk Aversion," Working Paper, University of London.
- Schneider, Paul, and Fabio Trojani, 2015, "(Almost) Model Free Recovery," Working Paper, Swiss Finance Institute.
- Schweri, Urs, 2010, "Is the Pricing Kernel U-Shaped?," Working Paper, University of Zurich.
- Shefrin, Hersh, 2008a, *A Behavioral Approach to Asset Pricing*, second Edition, Elsevier, Amsterdam.
- Shefrin, Hersh, 2008b, "Risk and Return in Behavioral SDF-Based Asset Pricing Models," *Journal of Investment Management*, 6, 1-18.
- Shimko, David C., 1993, "Bounds of Probability," *Risk Magazine*, 6, 33-37.

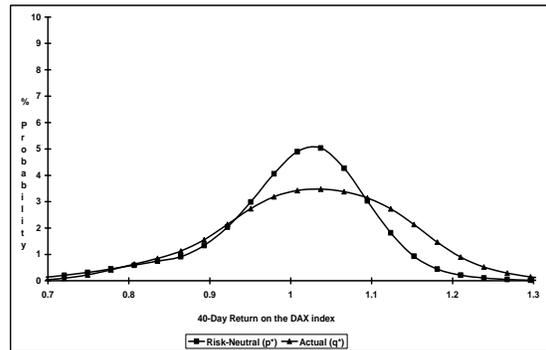
- Shive, Sophie, 2003, "On the Shape of the Option-implied Stochastic Discount Factor," Working paper, University of Michigan.
- Shive, Sophie, and Tyler Shumway, 2004, "A Non-decreasing Pricing Kernel: Evidence and Implications," Working Paper, University of Michigan
- Shive, Sophie, and Tyler Shumway, 2009, "Is the Pricing Kernel Monotonic?," Working paper, University of Notre Dame.
- Siddiqi, Hammad, and John Quiggin, 2016, "Differential Awareness and the Pricing Kernel Puzzle," Working paper, University of Queensland.
- Song, Zhaogang, 2012, "Expected VIX Option Returns," Working Paper, Federal Reserve Board.
- Song, Zhaogang, and Dacheng Xiu, 2015, "A Tale of Two Option Markets: State-Price Densities and Volatility Risk," *Journal of Econometrics*, forthcoming.
- Stutzer, Michael, 1996, "A Simple Nonparametric Approach to Derivative Security Valuation," *Journal of Finance*, 51, 1633-1652.
- Tversky, A., Kahneman, D., 1992, "Advances in Prospect Theory: Cumulative Representation of Uncertainty," *Journal of Risk and Uncertainty*, 5, 297-323.
- Veronesi, Pietro, 2004, "Belief-Dependent Utilities, Aversion to State-Uncertainty, and Asset Prices," Working Paper, University of Chicago.
- Wallmeier, Martin, 2015, "Mispricing of Index Options with Respect to Stochastic Dominance Bounds?," Working paper, University of Fribourg.
- Wang, M., M. O. Rieger, and T. Hens, 2013, "How Time Preferences Differ: Evidence from 52 Countries," Swiss Finance Institute Research Paper, No. 09-47.
- Weber, Thomas, 2006, "Estimating Risk Aversion in the European CDO Market," Working Paper, University of Konstanz, Germany.
- Wu, Liuren, 2006, "Dampened Power Law: Reconciling the Tail Behavior of Financial Security Returns," *Journal of Business*, 79, 1445-1473.
- Yaari, M.E., 1987, "The Dual Theory of Choice Under Risk," *Econometrica*, 55, 95-115.
- Yang, Jun, 2009, "Semiparametric Estimation of Asset Pricing Kernel," *Applied Financial Economics*, 19, 257-272.
- Zdorovenin, Vladimir, and Jacques Pezier, 2011, "Does Information Content of Option Prices Add Value for Asset Allocation?," Working Paper, University of Reading, UK.
- Ziegler, Alexandre, 2002, "State-Price Densities under Heterogeneous Beliefs, the Smile Effect, and Implied Risk Aversion," *European Economic Review*, 46, 1539-1557.
- Ziegler, Alexandre, 2007, "Why Does Implied Risk Aversion Smile?," *The Review of Financial Studies*, 20, 859-904.

Figure 1. Risk-Neutral and Actual Distributions for Four International Indices. The actual distributions are calculated with the same return horizon as the time-to-expiration of the options. For the US we used the historical sample from Sep 2, 1997 to Aug 15, 2003, for Germany from Jan 2, 1997 to Oct 9, 2003, for the UK from Jan 2, 1997 to Oct 9, 2003, and for Japan from Jan 5, 1998 to Oct 10, 2003. Returns are reported as 1 plus the rate of return.

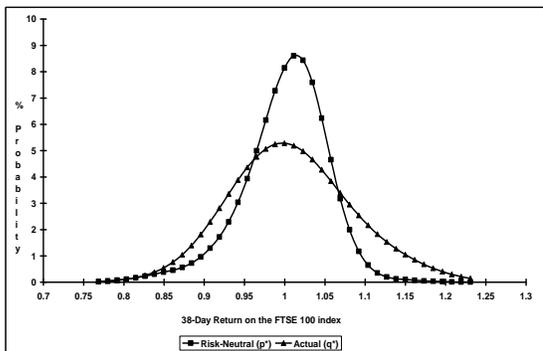
US  
45-Day options on Aug 15, 2003



Germany  
40-Day options on Oct 11, 2003



UK  
38-Day options on Oct 14, 2003



Japan  
41-Day options on Oct 11, 2003

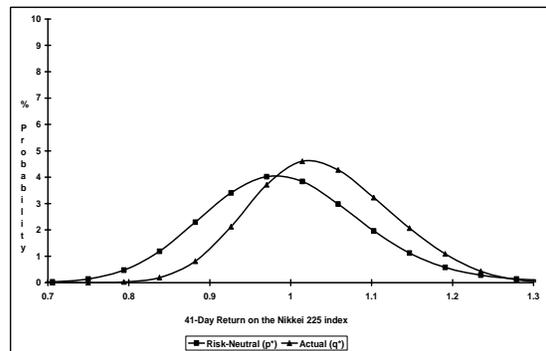
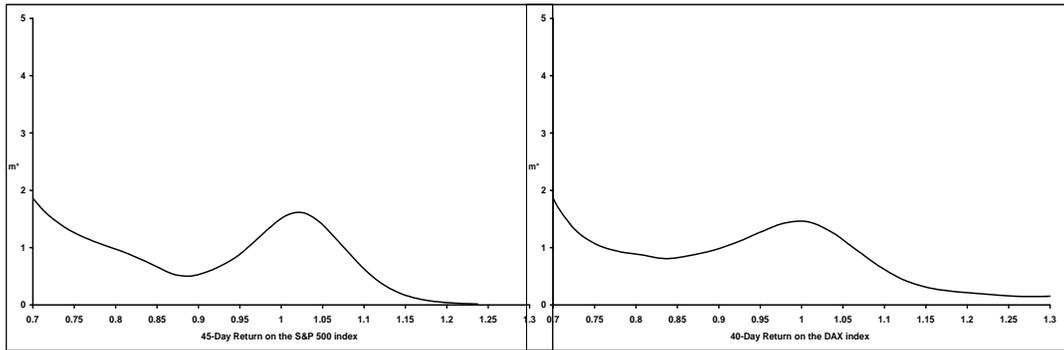


Figure 2. Empirical Pricing Kernels for Four International Indices. Typical post-1987 stock market crash implied pricing kernels. The pricing kernels are calculated as the ratio of the option implied risk-neutral distribution and the historical smoothed return distribution. Returns are reported as 1 plus the rate of return.

US  
45-Day options on Aug 15, 2003

Germany  
40-Day options on Oct 11, 2003



UK  
38-Day options on Oct 14, 2003

Japan  
41-Day options on Oct 11, 2003

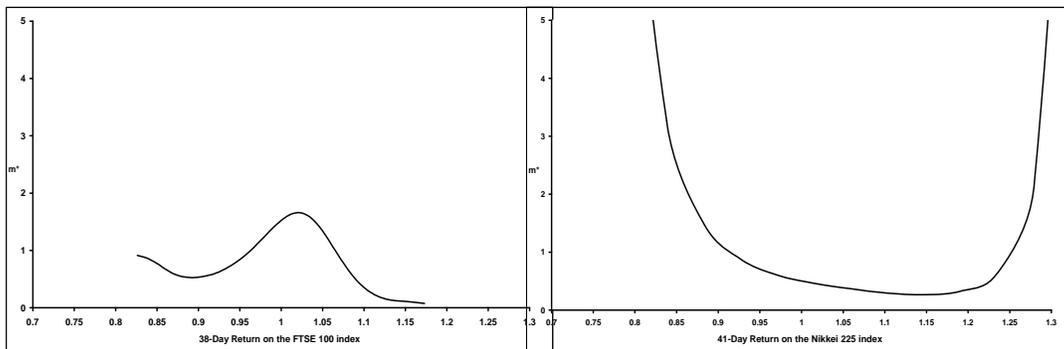
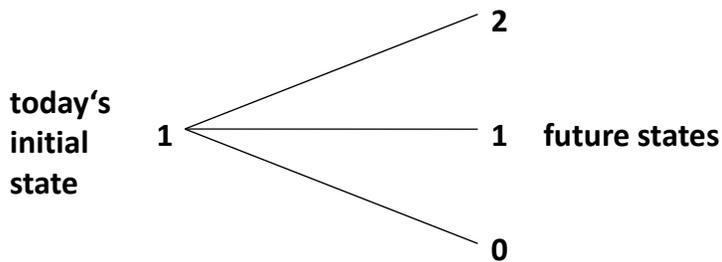


Figure 3. Risk-neutral Probabilities vs. Risk-neutral Transition Probabilities

In Panel A we depict the typical situation of a tree emanating from today's initial state (1) and moving to several future states (0, 1, and 2). In Panel B, we depict the data requirements of Ross (2015) where, in addition, one also needs to know the (hypothetical) transition probabilities from alternative states today (0 and 2) to all future states.

### Panel A: Risk-Neutral Density



### Panel B: Risk-Neutral Transition Probabilities

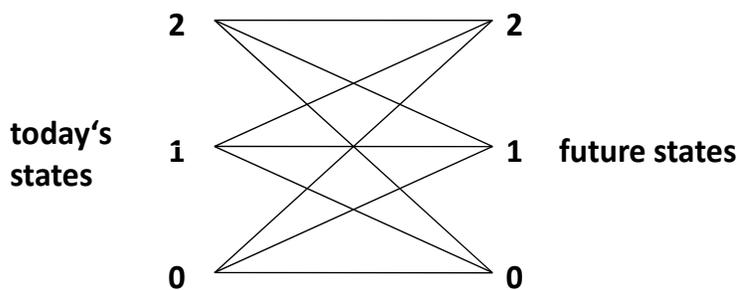


Figure 4. Time series of the location of local maxima of the empirical pricing kernel

The physical distributions are obtained by calibrating a GARCH(1,1) on a monthly basis using the past 3,500 daily returns. The risk-neutral distributions are obtained by applying the fast and stable method of Jackwerth (2004). The equity risk-premium is set to 5%. Each month we track the moneyness level at which the local minimum in the empirical pricing kernels occurs. Whenever the local minimum is at a moneyness level above 1.10 or whenever the pricing kernel is globally u-shaped, we assign a value of 1.10.

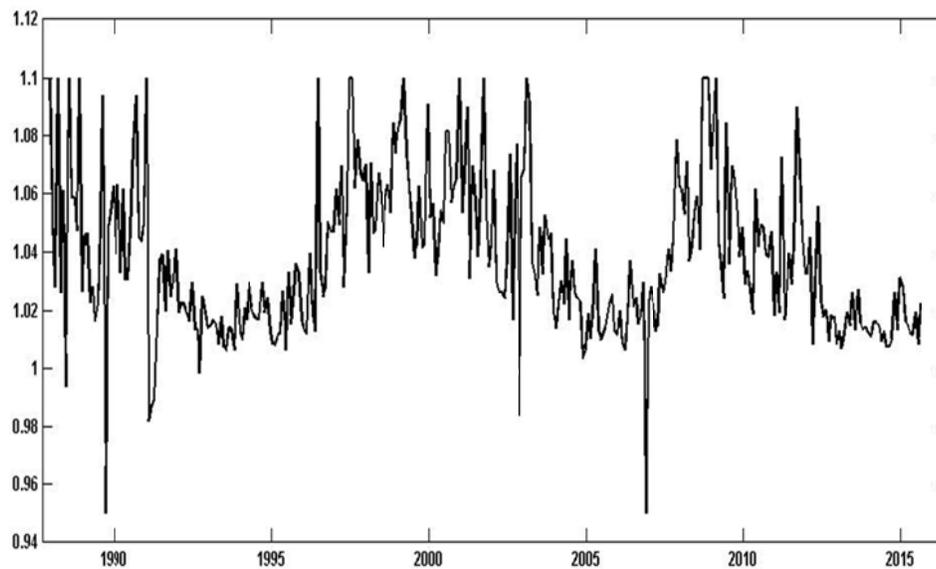


Figure 5. Friedman and Savage (1948) utility function

We depict a utility function along the lines of Friedman and Savage (1948) with concave-convex-concave segments in the return dimension. We also depict the concavified version of the utility function.

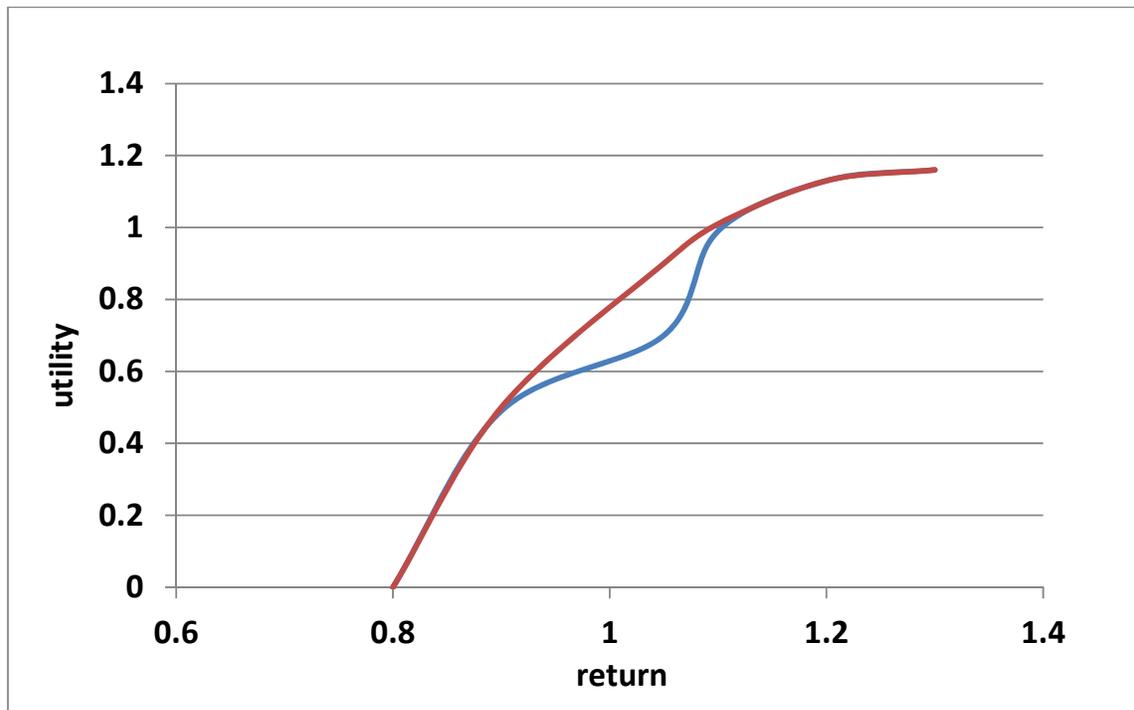


Figure 6. The pricing kernel with ambiguity over volatilities

The pricing kernel based on the Klibanoff, Marinacci, and Mukerji (2005) model with ambiguity over volatilities, projected onto returns. The reference pricing kernel puzzle is taken from figure 1, panel A, in Jackwerth (2004, p. 57).

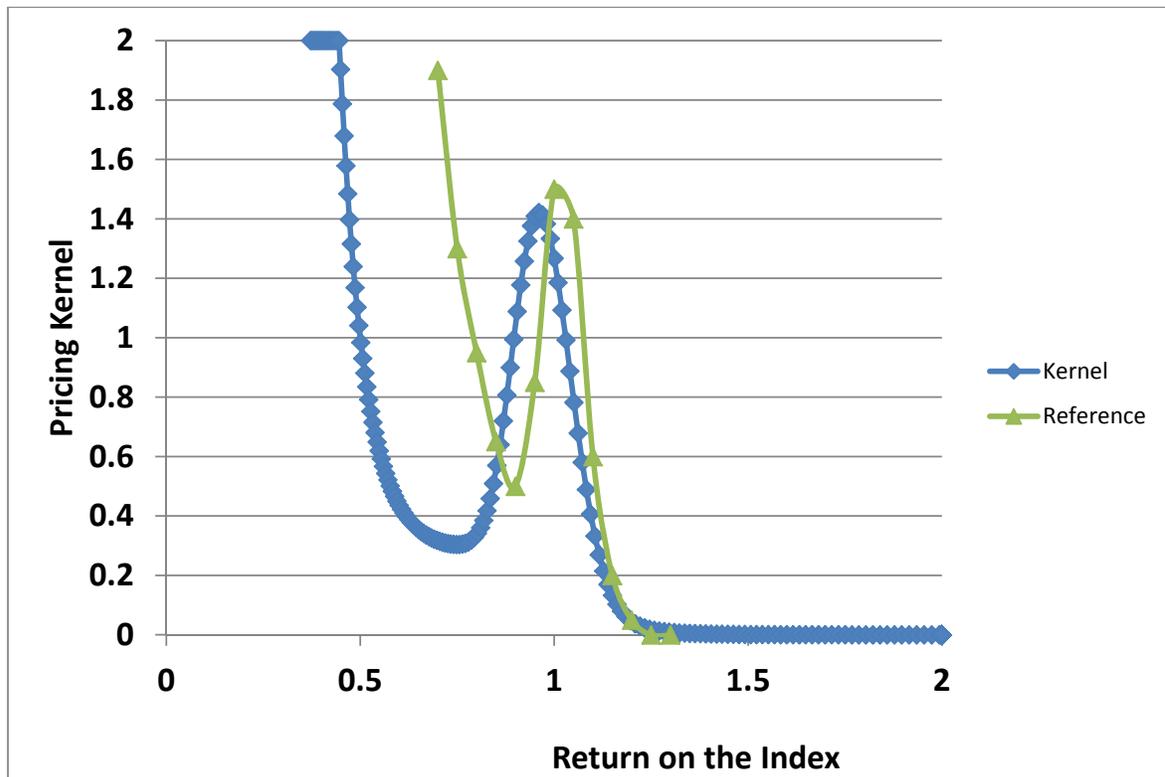


Figure 7. The pricing kernel with ambiguity over market crashes

The pricing kernel based on the Klibanoff, Marinacci, and Mukerji (2005) model with ambiguity over market crashes, projected onto returns.

